UK National Screening Committee

Rapid Review of Screening for Gestational Diabetes

External review against programme appraisal criteria for the UK National Screening Committee

Version: Final publication document

Author: Costello Medical

Date: June 2021

The UK National Screening Committee secretariat is hosted by Department of Health and Social Care.

About the UK National Screening Committee (UK NSC)

The UK NSC advises ministers and the NHS in the 4 UK countries about all aspects of population screening and supports implementation of screening programmes. Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK N S C recommendations.

UK National Screening Committee, Southside, 39 Victoria Street, London, SW1H 0EU

www.gov.uk/uknsc

Blog: https://nationalscreening.blog.gov.uk/

For queries relating to this document, please contact: https://view-health-screeningrecommendations.service.gov.uk/helpdesk/

Published September 2021

Contents

Plain English summary

Gestational diabetes mellitus (GDM) can happen in pregnancy when a woman cannot control her blood sugar levels. If not properly controlled, GDM can be dangerous for the mother and her baby. There is currently no population screening programme for GDM in the UK. However, the UK National Institute for Health and Care Excellence (NICE), recommends that women with risk factors are tested. Risk factors include obesity and family history of diabetes. Women are tested for GDM using a blood glucose test during pregnancy. Apart from NICE guidance, there are many different recommendations on how GDM should be diagnosed. There is no agreement on the best test.

This review looked to see if all pregnant women (not only those with risk factors) should be screened for GDM in the UK. It aimed to find evidence on:

- the risks of negative outcomes for the mother and baby that are related to increases in the mother's blood sugar. These are increases that are not high enough to cause a diagnosis of GDM according to the NICE test but are still above normal
- the best screening tests to find women at risk of GDM in pregnancy
- the best treatments for lowering blood sugar and stopping negative outcomes in women with GDM found by screening.

The main findings from the review were:

1. It is clear that increased blood sugar levels result in negative outcomes for the mother and baby. However, it is not clear what the cut-off value should be for a screening test to decide if a woman is 'at risk'.
2. There is no better test currently available than the glucose test used by NICE. This test is not fully accurate and may be risky. This is because the pregnant woman is given a sugar solution to drink when she might have poor blood sugar control.
3. The effects of treatments for GDM are unclear in women who have been found to be at risk through a screening test.

Based on the findings, a population screening programme for GDM in the UK is not

recommended. This topic will be reviewed again in 3 years' time.

Executive summary

Purpose of the review

This review was conducted to assess whether there is sufficient evidence to consider introducing a population screening programme for gestational diabetes in pregnant women.

Background

Gestational diabetes mellitus (GDM) is a condition characterised by elevated blood glucose and insulin resistance that is first detected during pregnancy. In healthy pregnancies, increased insulin resistance is a necessary physiological change that facilitates adequate carbohydrate supply for the fetus. But in pregnant women with GDM, hyperglycaemia and resistance to insulin is overly pronounced. GDM can develop during any stage of pregnancy, but most commonly presents in the second or third trimester.

GDM is diagnosed through assessment of maternal blood glucose levels, most often by fasting plasma glucose (FPG), an oral glucose challenge test (GCT) or oral glucose tolerance test (OGTT). However, there is an ongoing debate as to the specific type of test that should be used and the threshold at which GDM should be diagnosed. There is some agreement that the OGTT is the most appropriate test. However, there are issues with its administration and use as a screening test, as it involves glucose loading, which may be associated with side effects and could be harmful for women with impaired glucose tolerance, i.e. those that the test would most likely be used in. Agreement is less clear regarding the level of glucose in the loading solution (e.g. 75 g vs 100 g); length of time the test should be taken over (e.g. 1 vs 2 vs 3 hours); the thresholds at different timepoints at which a woman is diagnosed with GDM; as well as how many abnormal values should result in a GDM diagnosis. Due to the heterogeneity of glucose intolerance definitions and thresholds, some women with high glucose intolerance are not diagnosed with GDM.

Different recommendations for the tests and diagnostic thresholds are given by different national and international organisations. This translates into heterogeneity in studies of epidemiology and natural history of GDM, accuracy of screening tests and efficacy of interventions used to treat GDM. For example, while it is generally accepted that hyperglycaemia in pregnancy leads to adverse outcomes, the threshold at which the risk becomes relevant (i.e. how the at-risk group of pregnant women is defined) is unclear. Similarly, it is not known which test should be used to establish that threshold. In a study that compared the cost-effectiveness of screening women with and without risk factors for GDM, universal population screening (of all women

regardless of risk factors) was not found to be cost-effective compared with no screening or with risk-based screening.[1]

Once diagnosed with GDM, there exist effective treatments for GDM, including insulin, diet management and increased exercise. However, these have so far been studied in women tested for GDM due to having specific risk factors and it is unlear how effective they would be in women otherwise low risk but diagnosed with GDM through a population screening programme.

Focus of the review

This rapid review aimed to identify evidence published since the last UK NSC review, (based on the HTA report searches which were conducted in 2009), in answer to the following questions:

- Question 1: what are the risks of short and long-term adverse outcomes associated with incremental increases in maternal blood glucose level in the newborn?
- Question 2: what are the most effective screening tests or strategies to identify women at risk of hyperglycaemia in pregnancy or GDM?
- Question 3: what is the most effective intervention for lowering glucose levels in screen-detected pregnant women with GDM and preventing adverse perinatal outcomes?

Recommendation under review

Based on the 2010 UK NSC review of the evidence, screening for GDM in pregnant women is not currently recommended in the UK. However, women considered to be at high risk of GDM (based on risk factors) undergo testing with a 2-hour 75 g OGTT, based on guidance from NICE.[2]

Findings and gaps in the evidence of this review

The aim of question 1 was to identify associations between incremental increases in glucose levels that are elevated from normal in a low risk population (i.e. those not considered to be at risk of GDM according to NICE criteria or those treated for GDM) and the risks of adverse pregnancy and neonatal outcomes. This would allow for the characterisation of a 'low risk' population that may benefit from screening for GDM in those who are not currently covered by the NICE recommendation. For this question, moderate-to-high quality evidence for a wide number of pregnancy and neonatal outcomes was identified. The evidence was judged to be broadly applicable to the UK clinical setting. However, applicability to the review question was limited, as in most studies, the population of mild hyperglycaemia overlapped with women considered to

be at high risk of GDM, as covered by the NICE guideline. Although none of the studies selected women for specific risk factors, only 2 studies limited inclusion to low risk women with glucose thresholds indicative of mild hyperglycaemia who would not be diagnosed with GDM by the NICE guideline thresholds.

The review identified clear associations from a large volume of evidence between any elevated glucose and increased risk of several outcomes: C-section, induction of labour, macrosomia and large for gestational age (LGA). Macrosomia and LGA were also significantly increased in women who would not currently be identified as at risk by the NICE guideline, but neither C-section nor induction of labour was reported by either study investigating low risk women. Furthermore, a clear glucose threshold for increased risk could not be identified for any outcome, mostly due to the limited evidence on single thresholds. This is supported by the finding from previously published work that there is a continuum of risk across increasing glucose levels and no clear cut-off point.[3-5] On this basis, Criterion 1 was judged to be not met.

For question 2, despite the considerable size and reasonable quality of the evidence base, no screening test without glucose loading/challenge (with OGTT as the diagnostic reference standard) was found to be superior to using OGTT as a screening test on its own. In other words, none of the studies found a screening strategy that achieved test accuracies where both specificity and sensitivity were high enough to consider the test reliable and able to replace the current test used by NICE (2-hour 75 g OGTT), which involves glucose loading and therefore poses some risk of harm to women who are already suspected to be at risk of glucose intolerance. Using any of those strategies and only applying OGTT in screen-positive women would likely miss a considerable proportion of GDM (at a high threshold) or result in most women having to undergo OGTT anyway (at a lower threshold). Therefore, the best currently available test is the diagnostic OGTT test. This has drawbacks of uncertainty, because no better diagnostic test is available to compare against as a reference standard and women with GDM do not always show symptoms to make clinical diagnosis reliable. The OGTT test also carries a risk of harm of glucose loading and the consequences of using it in all pregnant women (including these at low risk) are unknown. Given the uncertainty around the accuracy and acceptability of the OGTT test (if used for screening) and lack of a better test, criterion 4 was judged to be not met.

For question 3, the aim was to identify the efficacy of interventions – compared with other interventions, no treatment or usual care – for lowering glucose levels and preventing adverse outcomes in pregnant women with screen-detected GDM (i.e. from a low-risk population that would not be identified by the current NICE pathway). However, due to unclear reporting on how GDM was detected, studies with populations with any GDM were included, in order to avoid limiting the available evidence. For populations with any GDM (i.e. not just screen-detected), there was a

moderate-to-high quality evidence base from 4 systematic literature reviews (SLRs; including 26 randomised controlled trials [RCTs]) and 8 RCTs in women with GDM treated with insulin, glibenclamide/glyburide, metformin or lifestyle interventions, such as diet or exercise. However, evidence was lacking in 2 key areas:

1) Comparison of interventions with placebo or usual care to allow evaluation of the benefit in treatment versus no treatment – the only comparisons were glybenclamide vs placebo in one RCT and lifestyle intervention vs usual care in 4 RCTs and 2 SLRs.
2) Studies including populations with screen-detected GDM to allow evaluation of the impact of treatment in a specific population where treatment is initiated in an early phase of the condition after identification by screening – only one RCT specified that GDM was screen-detected.

In the study comparing glibenclamide with placebo there was no evidence that treatment with glibenclamide significantly improved maternal or neonatal outcomes. In the studies comparing lifestyle interventions with usual care results were inconsistent. The only outcome for which risk was conclusively lower for dietary modification was pre-eclampsia, but this was only reported by 1 study. Risk was reported as lower with dietery modification for other outcomes including C-section, macrosomia, LGA, neonatal hypoglycaemia and shoulder dystocia by some studies, but in each case there was at least one other study that found no significant difference between dietary modification and usual care. Furthermore, in the one RCT that specified that GDM was screen-detected, there was no significant difference in any reported outcome (pre-eclampsia; gestational hypertension; C-section; LGA or neonatal intensive care unit [NICU] admission) between nutritional counselling and usual care. There is no certain evidence for the benefit of treating low-risk women with screen-detected GDM and therefore, criterion 9 was judged to be not met.

GDM and hyperglycaemia are important health problems. However, it is unclear whether benefits of treatment would outweigh the harms in low-risk women, if universal screening for GDM were to be introduced. This is because of uncertainties around the thresholds at which women should be considered at risk; the lack of a safe

and practical test or lack of data supporting the use of OGTT as a screening test; and lack of data supporting benefits from currently available interventions in screendetected women.

Recommendations on screening

Based on the evidence identified in this review, population screening for GDM is still not recommended. However, NICE guidelines should still be adhered to for women at high risk.

Limitations

Methodological limitations included limiting the searches to only including peerreviewed, English-language journal articles. The titles, abstracts and full texts were screened by 1 reviewer, with a second reviewer verifying all included, 10% of excluded decisions and any articles where there was uncertainty about their inclusion.

Evidence uncertainties

For question 1, evidence is lacking on the risk of adverse outcomes and the threshold at which these risks become significant for women who are currently not being tested for GDM in the UK – but who may have mild hyperglycaemia. In other words, it is unclear what is the population that should be defined as screen-positive if screening were to be introduced.

For question 2, the evidence indicates that no screening tests are superior to the currently used diagnostic test, the OGTT. The harms, especially to pregnant women who may have impaired glucose tolerance, as well as the acceptability of OGTT, are unclear.

For question 3, it is uncertain whether the conclusions of studies in women with clinically detected GDM also apply to women with screen-detected GDM.

Introduction and approach

Background

Clinical burden of disease

Gestational diabetes mellitus (GDM) is a condition characterised by elevated blood glucose and insulin resistance that is first detected during pregnancy. It can develop during any stage of pregnancy, but most commonly presents in the second or third trimester.[6] In healthy pregnancies, increased insulin resistance is a necessary physiological change that facilitates adequate carbohydrate supply for the fetus and the stimulation of fetal pancreatic insulin as an essential growth hormone, to meet the increased energy demands of pregnancy. However, in pregnant women with GDM, hyperglycaemia and resistance to insulin is overly pronounced.[7]

The most common form of GDM (~80% cases) is characterised by pancreatic β-cell dysfunction, where β-cells are no longer able to accurately detect blood glucose concentration or to adequately control release of insulin. This occurs following chronic insulin resistance, which is thought to occur in addition to the normal insulin resistance in pregnancy.[7] In addition, neurohormonal networks (e.g. leptin, adiponectin) along with several organ systems (e.g. pancreas, adipose, liver, muscle, gut, brain, placenta) may play a role in the pathogenesis of GDM.[7]

Evidence has consistently demonstrated that pregnant women with GDM and their newborns are at an increased risk of adverse perinatal outcomes, including large for gestational age (LGA), macrosomia, caesarean section and pre-eclampsia.[8] While GDM generally resolves after the baby is born, the effects of GDM on the mother and child may last beyond the timeframe of pregnancy, increasing the risk of longer-term maternal complications such as type 2 diabetes and cardiovascular disease,[9, 10] as well as the child's risk of obesity and associated cardiometabolic outcomes.[11, 12] Whether there exists a physiological link between maternal hyperglycaemia and perinatal and long-term adverse outcomes remains unclear; though it has been hypothesised to result from epigenetic remodelling due to intrauterine metabolic and inflammatory dysregulation.[13]

Prevalence of GDM

The estimated prevalence of GDM is influenced by several factors, including population characteristics, e.g. ethnicity or obesity; diagnostic criteria, e.g. glucose thresholds; and screening strategies, e.g. general or targeted screening.[4] As explored in more detail in the subsequent sections of this introduction, there is heterogeneity in

such factors, making it challenging to compare prevalence estimates at country-, regional- or even individual study-level. Another factor contributing to prevalence heterogeneity is consensus (or lack thereof) around the diagnostic criteria (discussed in more detail below), whereby estimates vary based on what test and thresholds have been applied within a study, making comparisons difficult.

An SLR conducted by Farrar (2016), which aimed to estimate the prevalence of GDM in the UK and Ireland, identified 13 studies on 16 cohorts of women. Reported prevalence across the studies varied substantially, ranging from 1.0% to 24.3%.[4, 14, 15] Prevalence estimates shifted from being consistently around 1 to 3% in studies carried out prior to 2010, which largely used the World Health Organization (WHO), [16, 17] diagnostic criteria, to a wider range of 8 to 24% for studies undertaken post-2010. This likely reflects the introduction of the 2010 International Association of the Diabetes and Pregnancy Study Groups (IADPSG) criteria that uses a lower fasting glucose threshold than WHO (Table 1), along with increasing trends in maternal older age and overweight/obesity during pregnancy.[4, 5] Use of IADPSG criteria resulting in a prevalence of GDM that was higher (35.5%) than the 10.6% found when using a procedure with higher glucose thresholds has been demonstrated.[18, 19]

The few studies that further stratified prevalence by population characteristics found that estimates generally increased with age, where prevalence of GDM was around 4 times higher in women over 40 years compared with women under 20 years old. GDM prevalence is also strongly correlated with ethnicity; specific ethnicities with higher prevalence are Hispanic, African, Native American, South or East Asian Pacific Islands or Indigenous Australians, whereas lowest GDM risk is found among women of Anglo-European descent.[20] The increase was up to 11 times higher in 1 study for South Asian women (prevalence of 4.4%) compared with White European women (prevalence 0.4%), [21] but was more consistently around 5 times higher in Asian women across other studies.[4] Other systematic reviews have investigated the impact of obesity on the prevalence of GDM, with 1 finding that there was a 0.9% increase in prevalence per 1 unit increase in body mass index (BMI). [22] One SLR found a strong correlative relationship between pre-pregnancy overweight/obesity and risk GDM, regardless of whether the assessment of BMI was self-reported, measured or take from hospital records.[23] A further analysis of the dose-response relationship found that GDM risk increased by 4% per 1 unit increase in BMI.[23]

In the Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) study, the overall prevalence of GDM across 15 international centres using IADPSG criteria averaged at

17.8%.[15] The evidence demonstrates that it is currently difficult to estimate the true prevalence of GDM with heterogeneity resulting from biological differences rather than being due to varied diagnostic criteria.

Screening for GDM

Diagnostic tests and thresholds

GDM is diagnosed through assessment of maternal blood glucose levels. This is most often performed by measuring fasting plasma glucose (FPG) and using an oral glucose tolerance test (OGTT), whereby blood glucose is first measured after an extended fast (≥8 hours e.g. overnight) before a glucose solution is ingested. Blood glucose concentration is then measured at 1 or more specific timepoints.[6] While the glucose levels of healthy pregnant women will be tightly controlled and quickly return to normal following glucose being metabolised, levels in women with GDM will remain higher for longer or even increase over a period of time.[24, 25] OGTTs typically involve 75 g or 100 g glucose loads with measurements taken at 1, 2 or 3 hours post-ingestion. A 50 g 1hour glucose challeng test (GCT) may also be used, particularly in the context of an initial test to stratify risk.[26]

Along with differences in the recommended glucose dose and time period for measuring plasma glucose post glucose ingestion, recommendations for the threshold value at which a diagnosis of GDM should be made (diagnostic threshold) and the required number of abnormal values have evolved over time and still vary considerably in current practice. Waugh (2010) suggests that assessment of FPG levels alone may be sufficient in the future once diagnostic thresholds are agreed, with advantages including lower resource use and avoidance of side effects, such as nausea and vomiting, associated with the ingestion of glucose-containing liquid.[5] At present, assessment of at least 1 OGTT value is recommended by clinical guidelines (Table 1).

Different recommendations for the tests and diagnostic thresholds as given by national and international organisations are summarised in Table 1. Early recommendations began with 3 sets of similar criteria published around the 1960–80s by O'Sullivan and Mahan,[27] later endorsed by the National Diabetes Data Group (NDDG)[28] and modified with the addition of a 50 g 1-hour GCT by Carpenter and Coustan.[29] Aside from the 50 g GCT and slight variations in the diagnostic threshold, all recommended using an initial FPG test, followed by a 100 g 3-hour OGTT, and required at least 2 values to be abnormal for a GDM diagnosis. The American Diabetes Association (ADA) also adopted the selective 2-step approach with an initial 50 g GCT preceding a full diagnostic 100 g 3-hour OGTT using the NDDG criteria.[30] Requiring 2 abnormal values for a diagnosis was also later adapted by the WHO in

several guidelines[14, 16, 17, 31] that recommend measuring FPG and/or a 75 g 2-hour OGTT, with the 2 abnormal glucose values emerging either from the same test or repeated tests. Notably, these were the same thresholds that were used for diagnosing type 2 diabetes mellitus, an approach that is no longer considered appropriate. 2010 saw the first major change to the status quo in GDM diagnosis when the IADPSG recommended a 1-step method and required just 1 plasma glucose value to exceed the threshold (either 5.1 mmol/L for FPG, 10.0 mmol/L for 1-hour 75 g OGTT or 8.5 mmol/L for 2-hour 75 g OGTT). This was based on the results of the landmark HAPO study, where cut-offs were calculated based on associations between maternal/neonatal complications and maternal glucose levels. [32] This approach may require fewer women to be screened in order to achieve a reduction in the number of adverse outcomes due to GDM. [18, 19] These criteria were later adopted by several other bodies including the ADA, [33] the WHO [34] and the Endocrine Society of the US [35] and are widely used in countries outside the US.[36] However, the 1-step approach remains contentious and is not globally accepted, largely due to concerns that the method results in overdiagnosis and overtreatment, with subsequent increases to cost and resource use without demonstrable improvement to maternal/neonatal outcomes.[19, 37]

The American College of Obstetricians and Gynecologists (ACOG),[38] and the National Institutes of Health (NIH)[36] in the US continue to recommend a selective 2-step approach, based on the thresholds originally recommended by NDDG [39] and Carpenter and Coustan [29] and an initial 50 g GCT before the full test, whilst the most recent ADA guidelines now recommend either a 1-step or 2-step approach.[26] In the UK there are further differences: the National Institute for Health and Care Excellence (NICE) recommend a conservative risk factor-based 1-step approach where GDM should be diagnosed if a woman has 1 or more risk factors (BMI >30 kg/m^2, previous macrosomic baby weighing ≥4.5 kg, previous GDM, family history of diabetes or higher risk ethnicity) and either FPG of ≥5.6mmol/L, or a 2-hour 75 g OGTT plasma glucose level of ≥7.8mmol/L.[2] This is similar to the 2010 Scottish Intercollegiate Guidelines Network (SIGN) criteria with slightly different glucose thresholds.[40] The 2 previous Health Technology Assessment (HTA) reports which assessed screening for gestational diabetes concluded that there was insufficient evidence to make a recommendation.[5, 41]

Table 1: Diagnostic criteria for the diagnosis of GDM

				N of	FPG		Glucose Threshold (mmol/L)			
		Load	**Organisation**	**Country**	**Year***	**Steps**	**Abnorma (g)**	1h	2h	3h	**Notes I Values**	OGTT	OGTT	OGTT
WHO[16, 17, 31, 34]	International	**2013**	1-step	1	5.1–6.9	75	10.0	8.5–11.0	-		The recommendation accepted the IADPSG 2010 thresholds but noted that the quality of evidence is "very low" and the strength of recommendation is "weak". It was also noted that diagnostic thresholds are likely arbitrary			
		1999	1-step	2	7.0	75	-	7.8	-		The FPG threshold was lowered to reflect the ADA 1997 recommendation for type 2 diabetes This recommendation used the same thresholds as those for detecting type 2 diabetes. This is no longer considered appropriate			
		1985	1-step	1 or 2	7.8	75	-	11.1	-		This recommendation used the same thresholds as those for detecting type 2 diabetes. This is no longer considered appropriate State that a single blood glucose value higher than the thresholds can establish the diagnosis, but that			
		1980	1-step	2	8.0	75	-	11.0	-		This recommendation used the same thresholds as those for detecting type 2 diabetes. This is no longer considered appropriate A clinical diagnosis was made on the basis of at least 2 abnormal values, either from the same test or repeated tests Values in the report were rounded to the nearest whole mmol/L			

NICE[2]	UK	**2015**	1-step riskfactor based		5.6	75	-	7.8	-	Recommend that screening only be conducted in women with any 1 of the following risk factor (BMI >30 kg/m^2, previous macrosomic baby weighing ≥4.5 kg, previous GDM, family history of diabetes
IADPSG[32]	International	**2010**	1-step	1	5.1	75	10.0	8.5	-	This recommendation was developed based on the results of the HAPO study and was the first to recommend a 1-step process requiring only 1 abnormal glucose value across the test Screening for Gestational Diabetes

	Organisation	**Country**	**Year***	**Steps**	**N of Abnormal Values**	**FPG**	**Glucose Threshold (mmol/L)**			**Load (g)**	**Notes**
							1h OGTT	2h OGTT	3h OGTT		
HTA[5, 41]	UK		**2010**	-	-	-	-	-	-	-	No additional recommendations on screening
			2002	-		-	-	-	-	-	Concluded that there was insufficient evidence to advocate for universal screening and that a highly selective policy based on maternal and pregnancy characteristics should be in place
SIGN[40]	Scotland		**2010** (updated 2017)	1-step with risk factors		5.1–6.9	75	-	8.6–11.0	-	At the first antenatal visit, women with risk factors should have HbA1c or FPG measured. At 24–28 weeks, all women with risk factors should have 75 g OGTT and women at low risk should have FPG
O'Sullivan and Mahan[27]	USA		**1964**	1-step	2	5.0	100	9.2	8.1	6.9	The thresholds were based on a glucose measurement in whole blood rather than plasma glucose and were first established to predict the postpartum development of type 2 diabetes
NDDG[28]	USA		**1979**	1-step	2	5.8	100	10.6	9.2	8.1	Suggested that a plasma glucose measurement would be preferred for simple testing
Carpenter-Coustan	USA		**1982**	2-step test should be given before the full	2 **criteria[29]**	- 5.3	50 100	7.2 10.0	- 8.6	- 7.8	Recommended that an initial 50 g screening diagnostic 100 g screening test
ACOG[38]	USA		**2013**	2-step		- 5.3 or 5.8	50 100	NR 10.0 or 10.6	- 8.6 or 9.2	- 7.8 or 8.0	Recommend the selective 2-step approach using either Carpenter-Coustan or NDDG threshold criteria for the 100 g OGTT

Endocrine Society of the US[35]	USA	**2013**	1-step	5.1	75	10.0	8.5	-	IADPSG recommendation adopted
NIH[36]	USA	**2013**	2-step	- 5.3 or 5.8	50 100	NR 10.0 or 10.6	- 8.6 or 9.2	- 7.8 or 8.0	Recommend the selective 2-step approach using either Carpenter-Coustan or NDDG threshold criteria for the 100 g OGTT The 15-member panel noted that resolution of the uncertainties regarding the 1-step approach would warrant possible revision to the current recommendations

Screening for Gestational Diabetes

Organisation	Country	Year*	N of Steps	Abnormal Values	FPG	Load (g)	Glucose Threshold (mmol/L) 1 h OGTT	2h OGTT	3h OGTT	Notes
ADA[26, 30, 33]	USA	**2014**	1-step		5.1	75	10.0	8.5	-	1-step approach described as "IADPSG consensus"
			or 2-step		- 5.3 or 5.8	50 100	7.8 10.0 or 10.6	- 8.6 or 9.2	- 7.8 or 8.0	Selective 2-step approach described as "NIH consensus" with an initial non-fasting 50 g 1-hour OGTT and uses CarpenterCoustan or NDDG criteria
		2011	1-step			5.1		8.5	-	Accepted the IADPSG recommendation
		1997	2-step		5.8	100	10.6	- 9.2	- 8.1	The 2-step approach includes an initial 50 g OGTT where a value of ≥7.8 mmol/L precedes the full diagnostic 100 g OGTT which uses the same thresholds as the NDDG criteria

Abbreviations: ACOG, American College of Obstetricians and Gynecologists; ADA, American Diabetes Association; BMI, body mass index; FPG, fasting plasma glucose; HAPO, Hyperglycemia and Adverse Pregnancy Outcome; HTA, Health Technology Assessment; IADPSG, International Association of the Diabetes and Pregnancy Study Groups; OGTT, oral glucose tolerance test; NDDG, National Diabetes Data Group; NICE, National Institute for Health and Care Excellence; NIH, National Institutes for Health; NR, not reported; SIGN, Scottish Intercollegiate Guidelines Network; UK, United Kingdome; USA, United States of America; WHO, World Health Organization

*The guidance is ordered by the year of the most recent published recommendations (year in **bold**) with all earlier recommendations from the same organisation listed below.

Aside from the aforementioned lack of reliability in estimating GDM prevalence, the inconsistency in the diagnostic criteria has numerous implications. Not least, it is challenging to
balance lower, more conservative test thresholds, which result in more women being diagnosed with GDM, with higher thresholds that limit the number of detected cases. It is important to consider the associated health and economic burdens such as unnecessary treatment in the former versus false reassurance and lack of treatment for women who may have GDM in the latter. Furthermore, a problem with the original O'Sullivan and Mahan criteria upon which all subsequent thresholds are based, is that the objective of these was to predict the risk of developing future type 2 diabetes or hyperglycaemia, rather than the risk of adverse maternal or neonatal outcomes.[37] In their 2013 guideline, the WHO also acknowledge that diagnostic thresholds are likely arbitrary, since the risk of adverse outcomes is continuous with increasing maternal blood glucose levels, as demonstrated in the HAPO study.[3]

Continuum of blood glucose values

While the association between GDM and pregnancy outcomes is well-established, the association between milder degrees of glucose intolerance during pregnancy and adverse pregnancy outcomes is less clear. There remains a need for better understanding of the relationship of varying levels of hyperglycaemia with maternal and neonatal outcomes, particularly the threshold at which risk is substantial enough to warrant intervention. Evidence on how milder gestational hyperglycaemia affects perinatal outcomes has emerged primarily from the HAPO study.[3] In response to discussions on the impact of inconsistency in diagnostic criteria for GDM, the study objectives were to determine predictive values for adverse pregnancy outcome by incremental glucose increase, in order to facilitate selection of internationally agreed diagnostic criteria.

During the study, 75 g 2-hour OGTT tests were administered to more than 25,000 nondiabetic pregnant women between 24- and 32-weeks' gestation, in 9 different countries. Data from HAPO mother-newborn pairs indicated an increase in the incidence of all 4 primary outcomes with increases in the 3 measures of blood glucose, both when these were analysed on a continuous scale and when categorised into 5 mg/dL increments.[3] After adjusting for confounders including age, BMI and family history of diabetes, an increase in 1 standard deviation of fasting blood glucose (FBG, 6.9 mg/dL) was associated with significantly increased odds of macrosomia (odds ratio [OR]=1.38; 95% confidence intervals [CI]: 1.32 to 1.44), primary caesarean section (OR=1.11; 95% CI: 1.06–1.15) and neonatal hyperinsulinemia (OR=1.55; 95% CI: 1.47 to 1.64), though the association with neonatal hyperglycaemia was not found to be significant.[3] However, no clear threshold in blood glucose values was observed that

could help to inform objective outcome-based diagnostic criteria for GDM. Instead, the results of the HAPO study suggest that GDM is not a clearly differentiated disease state and that the perinatal and long-term risks for mother and child should be considered within a continuum of glycaemic values.[42] This has far reaching implications for screening for GDM; without a threshold that can be used to define the at-risk group(s), screening will not be able to identify women in whom an intervention would be beneficial. Too high a threshold could cause women to miss out on treatment and too low a threshold could lead to overtreatment.

There is evidence from interventional studies (Australian Carbohydrate Intolerance Study in Pregnant Women [ACHOIS] and Landon 2009) indicating a benefit of diet and insulin as needed on the incidence of shoulder dystocia and measures of birth weight and size, in women with blood glucose lower than values traditionally considered to be clinically significant in diagnostic testing (based on investigator-defined fasting glucose and OGTT values)[43, 44] However, evidence on the specific maternal blood glucose levels at which treatment is warranted to prevent or minimise other adverse neonatal outcomes remains to be evaluated. In order to inform recommendations of implementing a screening programme, further data is needed on the levels of blood glucose at which lifestyle and pharmaceutical interventions are clinically and economically beneficial at reducing the risk of adverse outcomes for mothers and their newborns.[5]

Screening modalities

In addition to the lack of consensus on maternal blood glucose thresholds for diagnosis of GDM, approaches for how pregnant women should be selected for GDM screening also vary globally. In general, screening for GDM is usually performed at 24 to 28 weeks' gestation in order to coincide with the rise of insulin resistance that typically occurs during the second trimester. Women who do not have the ability to produce sufficient insulin to adapt to this resistance will consequently present with higher plasma glucose levels.[19]

In the UK, universal screening for GDM is not currently recommended; only women considered to be at high risk of GDM undergo testing based on guidance from NICE.[2] Indications for testing include BMI >30kg/m^2, previous macrosomic baby weighing ≥4.5 kg, previous GDM, family history of diabetes or minority ethnic family origin with a high prevalence of diabetes. Women who are considered at risk of GDM should be offered a
75 g 2-hour OGTT at 24 to 28 weeks' gestation. NICE guidance also recommends that pregnant women with a history of GDM should be offered early self-monitoring of blood

glucose, or a 75 g 2-hour OGTT as soon as possible after the first antenatal appointment in the first or second trimester, followed by a 75 g 2-hour OGTT at 24 to 28 weeks if the results of the first OGTT were normal.[2] The NICE 2015 guidance approach was found to be cost-effective when compared with the 2013 WHO recommendations or 'no screening' from an NHS perspective. The same study also concluded that universal screening (application of diagnostic thresholds to all women regardless of risk factors) was not cost-effective against no screening or guidelinedirected risk-based screening.[1]

Current ADA guidelines also recommend selective screening of women considered to be at risk of GDM, based on the same risk factors used to select non-pregnant adults for Type 2 diabetes screening. These risk factors are comparable to those defined by NICE guidelines but also include women who are overweight as well as obese women (BMI ≥25 kg/m^2) and have additional risk including polycystic ovary syndrome, high cholesterol or haemoglobin A1c (HbA1c) levels, and a history of cardiovascular disease or hypertension.[26]

The WHO further suggest that glycosuria on dipstick testing (2+ or above on 1 occasion, or 1+ on two or more occasions) may be indicative of undiagnosed GDM, prompting a need for further diagnostic testing.[34] The benefit of screening for GDM in pregnant women without known risk factors is also currently being discussed, given the subsequent implications for treatment decision-making.

A number of organisations have recommended and/or implemented universal screening for GDM, including ACOG and the United States Preventative Services Task Force (USPSTF). In 2014, the USPSTF recommended screening for GDM in asymptomatic pregnant women after 24 weeks of gestation but stated that there was insufficient evidence on the benefits or harms of screening in low-risk women before 24 weeks' gestation. This guidance is currently under review.[45] The IADPSG (2010) similarly recommend that all women without known diabetes before pregnancy should undergo a 2-hour 75 g OGTT test at 24 to 28 weeks gestation.

In conclusion, there is still an ongoing discussion regarding the criteria that should be used for diagnosis of GDM as well as the population that should be defined as at-risk, and as such, current screening recommendations and diagnostic guidance have still not yet been agreed upon. Consensus on diagnosis criteria would allow for improved estimates of GDM prevalence, harmonisation in evidence generation and collection, and facilitate optimal treatment decision-making, in order to improve maternal and neonatal outcomes in GDM.

Treatment for GDM

The current NICE clinical practice guideline (NG3, 2015) recommends that women who have been diagnosed with GDM are referred to a dietitican, advised to eat a healthier diet and exercise regularly. For women with a FPG of <7 mmol/L at diagnosis, a trial of diet and exercise changes may initially be suggested but if blood glucose targets are not met with 1 to 2 weeks, metformin, insulin or metformin and insulin should be offered. For women with a FPG of ≥7 mmol/L at diagnosis, immediate treatment with insulin, with or without metformin is recommended, along with diet and exercise changes.[2] Similarly, guidance from the ADA states that lifestyle change is an "essential component" of GDM management and that additional medications (preferentially insulin as first-line) should be added if needed.[46] A recent appraisal of 14 guidelines from international organisations (including NICE and ADA) found commonalities across all guidelines. The main principles included lifestyle intervention, particularly nutrition therapy, as essential; use of medical therapy if needed to achieve glycaemic targets; regular self-monitoring of blood glucose. A main difference was the preferred agent for medical therapy to treat hyperglycaemia. In 6 guidelines, insulin was recommended as first-line therapy, whereas in another 6, oral antidiabetic agents (for example, metformin) were recommended.[47]

At present, treatment is only recommended for women diagnosed with GDM based on the NICE pathway. There are no recommendations for how to treat low-risk women who would be diagnosed with GDM should a population screening programme be introduced.

Current policy context and previous reviews

Screening for GDM in pregnant women is currently not recommended in the UK. The initial UK NSC recommendation not to introduce a GDM screening programme was based on a 2002 HTA report which concluded that screening for GDM did not meet sufficient UK NSC criteria.[41] A precise definition of GDM was lacking and adverse outcomes of increased glucose levels were reported mostly as macrosomia, the thresholds for which were considered somewhat arbitrary and not distinguishing between larger babies and those with abnormal growth, where treatment may be beneficial. No standardised test to screen for GDM was available and there was a concern that some women with low levels of glucose intolerance and who are not at risk of adverse outcomes may suffer anxiety and inconvenience due to receiving the diagnosis.

This was followed by another HTA, in 2010, which incorporated the findings of the

HAPO and ACHOIS studies, and despite finding an increased knowledge base around the condition, there was still insufficient evidence to determine blood glucose levels at which interventions may provide benefit.[5] Currently, the risk-factor based testing is recommended by NICE, but it is unclear whether women without the NICE-specified risk factors could be at risk of adverse outcomes if their blood glucose values are elevated but not yet reaching the 7.8 mmol/L threshold specified by NICE.

This rapid review aims to identify evidence published since the last HTA report searches which were conducted in 2009, in answer to the following questions:

- what are the risks of short and long-term adverse outcomes associated with incremental increases in maternal blood glucose level in the newborn?
- what are the most effective screening tests or strategies to identify women at risk of hyperglycaemia in pregnancy or GDM?
- what is the most effective intervention for lowering glucose levels in screen-detected pregnant women with GDM and preventing adverse perinatal outcomes?

Objectives

This review aims to assess whether there is sufficient evidence to consider introducing a screening programme for GDM in pregnancy. Specifically, the review will focus on introducing a screening programme in the context of current recommendations by the NICE NG3 guideline, whereby testing for GDM by 75 g 2-hour OGTT is recommended for women with risk factors and diagnosis is made if the FBG value is ≥5.6 mmol/L or the 2-hour OGTT value is ≥7.8 mmol/L.

The review will appraise evidence on the questions in Table 2, which each relate to the criteria set out by the UK NSC for assessing the suitability of a screening programme.

Table 2. Key questions for the evidence summary, and relationship to UK NSC screening criteria

Criterion	Key questions	Studies Included
THE CONDITION		

1	The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including	What are the risks of short and long-term adverse outcomes in the newborn associated with	23 publications on 18 studies
	Criterion	**Key questions**	**Studies Included**
	development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease.	incremental increases in maternal glucose level?	
	THE TEST		
4	There should be a simple, safe, precise and validated screening test.	What are the most effective screening tests or strategies to identify women at risk of hyperglycaemia in pregnancy or GDM?	18 publications on 14 studies
	THE INTERVENTION		
9	There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered.	What is the most effective intervention for lowering glucose levels in screened detected pregnant women with GDM and preventing adverse perinatal outcomes?	17 publications on 12 studies

Methods

The current review was conducted by Costello Medical, in keeping with the UK National Screening Committee evidence review process. Database searches were conducted on 21 August 2019 to identify studies relevant to the questions detailed in Table 2.

Eligibility for inclusion in the review

Eligibility criteria for each question are presented in Table 3, Table 4 and Table 5 below.

Systematic literature reviews (SLRs) and meta-analyses (MAs) were considered for inclusion for all questions in this review. If the scope of an SLR or MA was closely aligned to 1 of the topics of this review, it was included in its own right. However, if the scope was not closely aligned to 1 of the topics of this review, but some of the included articles may have been of interest, the reference list of the SLR or MA was handsearched. Any relevant primary research articles identified that were relevant to this review were then included, but the SLR or MA itself was excluded.

Review process

The following review process was followed:

4. Each abstract was reviewed against the inclusion/exclusion criteria by 1 reviewer. Where the applicability of the inclusion criteria was unclear, the article was included at this stage in order to ensure that all potentially relevant studies were captured. A second independent reviewer provided input in cases of uncertainty, and validated all of the first reviewer's inclusions and 10% of exclusions. Any disagreements were resolved by discussion until a consensus was met.
5. Full-text articles required for the full-text review stage were acquired online if freely available, or through the Cambridge University Library. Any paywalled articles unavailable at the Cambridge University Library which were deemed to have high potential of being relevant to the review questions were purchased in consultation with the UK NSC.
6. Each full-text article was then reviewed against the inclusion/exclusion criteria by 1 reviewer, who determined whether the article was relevant to 1 or more of the review questions. A second independent reviewer provided input in cases of uncertainty and validated all of the first reviewer's inclusions and 10% of exclusions. Any disagreements were resolved by discussion until a consensus was met.

Table 3. Inclusion and exclusion criteria for review question 1 (Q1)

Domain	Population	Intervention	Comparator	Outcome	Study type	Setting	Other considerations
Inclusion criteria	Unselected pregnant women without preexisting diabetes or diagnosed GDM	**Prognostic factor/exposure** Elevated maternal glucose (identified by tests to detect GDM)	Maternal glucose defined as normal	Risks of adverse neonatal outcomes, including but not limited to: Perinatal mortality Mode of birth (including induction of labour) Macrosomia and LGA Birth injury (e.g. dystocia, brachial plexus neuropathy) Hypoglycaemia Admission to neonatal care unit Long-term neonatal outcomes (e.g. greater adiposity and cardiometabolic illhealth)	RCTs (control arms only), metaanalyses and systematic reviews, crosssectional studies, cohort studies	Tier 1: Studies conducted in the UK Tier 2: Studies conducted in highincome countries where the screening methods and technology are expected to be similar to that of the UK (OECD and EEA countries excluding South Korea and Mexico)*	Articles published in the English language since January 2009
Exclusion criteria	Women who are not pregnant Cohorts selected for the presence of a specific condition e.g. women with preexisting diabetes or GDM, women receiving treatment for diagnosed GDM, women selected for other risk factors Multiple pregnancies only	Any other prognostic factors if maternal glucose is not included	Any other comparators	Any other outcomes	Any other study design, including reports, international case series, narrative reviews, eligible countries editorials, are not presented commentaries, separately to letters, outcomes from conference abstracts or	Studies in ineligible countries, or case studies where outcomes for ineligible countries o t	Studies with full text not in the English language Studies published pre2009

h e r p u b l i c a t i o n t y p e s t h							a v e n o t b e e n p e e r r e v i e w e
a t h							d

Abbreviations: EEA, European Economic Area; GDM: gestational diabetes mellitus; LGA, large for gestational age; OECD, Organisation for Economic Co-ordination and Development; RCT, randomised controlled trial.

Where possible, risk was stratified by pregnancy characteristics (e.g. age, BMI, ethnicity). In the first instance, this review focused on evidence related to the UK population.

*A decision rule was formulated for Tier 2 evidence, depending on the level of available Tier 1 evidence for specific outcomes.

Table 4. Inclusion and exclusion criteria for review question 2 (Q2)

Domain	Population	Intervention	Comparator	Outcome	Study type	Setting	Other considerations

Inclusion criteria	Unselected pregnant women without preexisting diabetes	**Index test** Screening test to identify GDM, including but not limited to: • (O)GTT • (O)GCT • IGT • Fasting glucose • Maternal history or risk factors **Reference standard** The reference standard as defined by the study	No screening or current practice	Test accuracy, including but not limited to: • Sensitivity • Specificity • PPV • NPV • LR (AU)ROC curve	Diagnostic test accuracy studies, cross-sectional studies, cohort studies, metaanalyses and systematic reviews	<u>Tier 1:</u> Studies conducted in the UK <u>Tier 2:</u> Studies conducted in highincome countries where the screening methods and technology are expected to be similar to that of the UK (OECD and EEA countries excluding South Korea and Mexico)*	Articles published in the English language since January 2009

Exclusion criteria	Women who are not pregnant Cohorts selected for the presence of a specific condition e.g. women with preexisting diabetes, women receiving treatment for diagnosed GDM, women selected for other risk factors Multiple pregnancies only	Irrelevant index test or reference standard	Any other comparators	Any other outcomes	Any other study design, including RCTs, case reports, case series, narrative reviews, editorials, commentaries, letters, conference abstracts or other publication types that have not been peerreviewed	Studies in ineligible countries, or international studies where outcomes for eligible countries are not presented separately to outcomes from ineligible countries	Studies with full text not in the English language Studies published pre2009

Abbreviations: (AU)ROC, (area under) receiver operating characteristic; EEA, European Economic Area; GDM: gestational diabetes mellitus; OECD, Organisation for Economic Coordination and Development; (O)GCT, (oral) glucose challenge test; (O)GTT, (oral) glucose tolerance test; NPV, negative predictive value; PPV, positive predictive value; RCT, randomised controlled trial.

*A decision rule was formulated for Tier 2 evidence, depending on the level of available Tier 1 evidence for specific test parameters.

		• Insulin			reviews, quasiexperimental		

Table 5. Inclusion and exclusion criteria for review question 3 (Q3)

Domain	Population	Intervention	Comparator	Outcome	Study type	Setting	Other considerations
Inclusion criteria	Pregnant women with GDM	Pharmacological interventions, including but not limited to:	No intervention or relevant intervention	Pregnancy, maternal and neonatal a outcomes, including but not limited to:	RCTs, metaanalyses and systematic	<u>Tier 1:</u> Studies conducted in the UK	Articles published in the English language

	Babies born to mothers with GDM (age cut-off of 28 days)	• Metformin Non-pharmacological interventions, including but not limited to: • Gentle exercise • Diet Any combination of the above	Pregnancy outcomes: • Perinatal mortality • Mode of birth (including induction of labour) • Maternal gestational weight gain Maternal outcomes: • Maternal wellbeing • Postpartum haemorrhage • Method of infant feeding • Post-pregnancy type 2 diabetes Neonatal outcomes: • Macrosomia • Birth injury • Hypoglycaemia • Admission to neonatal care unit	studies, cohort studies, casecontrol studies	<u>Tier 2:</u> Studies conducted in highincome countries where the screening methods and technology are expected to be similar to that of the UK (OECD and EEA countries excluding South Korea and Mexico)*	since January 2009
Exclusion criteria	Women who are not pregnant Pregnant women without GDM Healthy newborns					

Any other interventions	Any other comparators	Any other outcomes	Any other study design, including case reports, case series, narrative reviews, editorials, commentaries, letters, conference abstracts or other publication	Studies in ineligible countries, or international studies where outcomes for eligible countries are not presented separately to ineligible countries	types that have not been peerreviewed	Studies published pre-2009 Studies with full text not in the English language

Abbreviations: EEA, European Economic Area; GDM: gestational diabetes mellitus; OECD, Organisation for Economic Co-ordination and Development; NPV, negative predictive value; PPV, positive predictive value; RCT, randomised controlled trial.

*A decision rule was formulated for Tier 2 evidence, depending on the level of available Tier 1 evidence for specific outcomes.

Appraisal for quality/risk of bias tool

Quality assessments were performed by 1 reviewer for each included study and independently verified by a second individual. Any discrepancies were discussed until a consensus was reached; if necessary, a third independent reviewer made the final decision.

The following tools were used to assess the quality and risk of bias of each study included in the review:

- Diagnostic accuracy of screening test studies: Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool[48]
- Accuracy of diagnostic model studies: Prediction model Risk Of Bias Assessment Tool (PROBAST)[49]
- RCTs: Cochrane Collaboration's "Risk of Bias" tool[50]
- Non-randomised interventional studies and observational studies: Risk of Bias in Non-randomised Studies – of Interventions (ROBINS-I) tool[51]

The full guidance used for the quality assessments is available in Appendix 4.

Databases/sources searched

The following databases were searched:

- MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print • Embase
- The Cochrane Library, including the following:
- Cochrane Database of Systematic Reviews (CDSR)
- Cochrane Central Register of Controlled Trials (CENTRAL) • Database of Abstracts of Reviews of Effects (DARE)

MEDLINE and Embase were searched simultaneously via the Ovid SP platform. The Cochrane Library databases were searched via the Wiley Online platform and DARE was searched via the Centre for Reviews and Dissemination (CRD) website.

Searches were run on 21 August 2019. Full details of the searches, including the search strategy for each database, are presented in Appendix 1.

Searches for Q1 were based on an adapted search strategy of Farrar 2016.[4] Adaptations included limiting to RCT, non-RCT and observational study designs using a well validated search filter[52] and addition of some exclusion terms, such as to exclude conference abstracts. Farrar 2016 did not include search terms for large for gestational age (LGA) or infant mortality; whilst LGA was included as an outcome in the SLR,

evidence on perinatal mortality was not included. Therefore, additional terms have been added for these outcomes, with the date limit for perinatal mortality terms altered to 2009 to capture studies reporting on this outcome that were not included in the Farrar 2016 SLR.

Searches for Q2 were also based on Farrar 2016 that included studies on tests based on glucose tolerance, and maternal history and risk factors; new evidence on the accuracy of those tests was date limited to October 2014.[4] Searches for studies on tests based on maternal screening and biomarkers not included in Farrar 2016 were date limited to 2009.

For Q3, the 3 SLRs that formed the evidence base only included RCT evidence. As such, only the search results identified through the RCT search filter were date limited to 2016. Non-RCTs and observational studies were date limited to 2009.

Question level synthesis

Criterion 1 — The condition should be an important health problem as judged by its frequency and/or severity

Question 1 – What are the risks of short and long-term adverse outcomes in the newborn associated with incremental increases in maternal glucose level?

The National Institute for Health and Care Excellence (NICE) guidelines currently recommend a 1-step risk factor based approach to screening for gestational diabetes mellitus (GDM), whereby women with a risk factor(s) (body mass index [BMI] >30 kg/m^2; prior macrosomic baby ≥4.5 kg; previous GDM; family history of diabetes; high risk ethnicity [South Asian, black Caribbean, Middle Eastern]) should undergo screening. GDM is diagnosed if a woman has either a fasting plasma glucose (FPG) of ≥5.6 mmol/L or a 2hour 75 g oral glucose tolerance test (OGTT) of ≥7.8 mmol/L.[2]

The aim of this question was to identify associations between incremental increases in glucose levels that are elevated from normal in a low risk population (i.e. those not considered to have GDM according to NICE criteria or those treated for GDM) and the risks of adverse pregnancy and neonatal outcomes. This would allow the characterisation of a "low risk" population that may benefit from screening for GDM who are not currently included in the NICE recommendation. This question was partly considered in the 'Screening' chapter of the rapid evidence synthesis on screening for hyperglycaemia in pregnancy, conducted by Waugh *et al*. in 2010 for the UK NSC, which highlighted the need to identify a glucose threshold at which women should be classified as being at high risk.[5] However, the specific risks of specific outcomes associated with elevated maternal glucose were not quantified.

This evidence synthesis includes a large systematic literature review (SLR), Farrar 2016, whose searches were conducted in October 2014 and were updated as part of this review.

Eligibility for inclusion in the review

This review searched for control arms of randomised controlled trials (RCTs), crosssectional and cohort studies, and SLRs and meta-analyses (MAs) of these study types, published since January 2009. Studies were included if the population comprised unselected pregnant women without pre-existing diabetes or other specific risk factors (i.e. not a population that would be eligible for screening as per the NICE definition of GDM). Eligible women had singleton pregnancies and had undergone assessment of glucose tolerance. Studies or data from subgroups of women treated for GDM were not eligible for inclusion. The prognostic factor of interest was elevated maternal glucose identified by diagnostic tests for GDM,

compared with normal maternal glucose levels. Studies were required to compare at least 1 category of elevated maternal glucose (as defined by the individual study) with NGT to allow identification of differences between women with elevated glucose (but not considered to have GDM) compared with NGT. Outcomes of interest for question 1 were risk of adverse pregnancy, neonatal or long-term offspring outcomes. Outcome data were reported as number of events, odds or risk ratios relative to glucose categories or unit increments in glucose. Studies that only reported on correlations were excluded as these would not allow characterisation of a specific cut-off threshold for elevated risk. Studies were restricted geographically to Organisation for Economic Cooperatione and Development (OECD) or European Economic Area (EEA) countries, excluding Mexico and South Korea.

A large SLR and MA conducted as part of a Health Technology Assessment (HTA) by Farrar and colleagues (2016), was identified and updated as part of the evidence synthesis for question 1. The aim of Farrar 2016 was to determine the association between graded increases in glucose level and risk of perinatal and longer-term outcomes and the eligibility criteria were largely aligned with the eligibility criteria of this rapid review for question 1. However, there were some differences. For example, the search strategy used in Farrar 2016 did not include terms for perinatal mortality or large-for-gestational-age (LGA), which was accounted for by adding search terms for these outcomes. Furthermore, Farrar 2016 did not use a geographic limit and included 7 studies from non-OECD/EEA countries (Singapore, South Korea, Pakistan, Iran and China). Separate results excluding these countries were not available from the MA, therefore this should be noted as a limitation when considering the generalisability of the results to a UK setting.

Description of the evidence

Three publications reported on the Farrar 2016 SLR. Seventeen primary publications from database searches were judged to be relevant to question 1. Two additional publications from database searches were reference linked to Farrar 2016 as they reported novel data on studies included in Farrar 2016. Ultimately, there were 22 publications on 18 unique studies (1 SLR and 17 primary research). The key details of the included studies are presented in Table 6. Figure 1 (Appendix 1) contains a full PRISMA diagram.

Farrar 2016 SLR and MA

Farrar 2016 included 57 studies in the qualitative synthesis and 37 studies in the MA. Key studies included were HAPO, Born in Bradford (BiB) and ATLANTIC-DIP (Diabetes in Pregnancy). Where publications reported on the same study or cohort, data from the most recent and comprehensive publication for each outcome was used. As required by the eligibility criteria, all studies used at least 1 of the 50 g glucose challenge test (GCT),

75 g OGTT or 100 g OGTT to assess glucose tolerance. The Farrar 2016 HTA also reported on a separate analysis of individual patient data (IPD) from the BiB birth cohort study. This is not considered separately in the discussion of the results because the BiB study was also included in the MA; however, full details of the outcomes from the IPD analysis are presented in Appendix 2 Table 47.

The results from the MA conducted as part of Farrar 2016 are reported as odds ratios (ORs) for specific outcomes per 1 mmol/L increment in glucose (i.e. a dose-response). Different estimates were produced based on glucose levels measured by the 1 h 50 g GCT, 75 g OGTT (FPG, 1 h and 2 h) and 100 g OGTT (FPG, 1 h and 2 h). In order to increase the number of studies and participants included in the comparisons, the results for the 75 g and 100 g OGTT were combined, with the assumption that the association between outcomes and increase in glucose were the same for both tests. The combined 75 g/100 g OGTT results are the ones discussed in the results of this rapid review, while full details of all outcomes for individual glucose tests are presented in Appendix 2 Table 47.

One additional publication on the HAPO study (Belfast site) and 1 additional publication on the ATLANTIC-DIP cohort were identified and included as supplementary to the Farrar 2016 SLR.[53, 54]

Other included studies from database searches

Of the 17 publications included as distinct from the Farrar 2016 SLR, 1 was a secondary analysis of an RCT, 4 were prospective cohorts and 12 were retrospective cohorts. Studies were conducted in the US (n=6), Turkey (n=3), Spain (n=2), Australia (n=1), Canada (n=1), England (n=1), Italy (n=1), Sweden (n=1) and Japan (n=1). All studies measured glucose using 1 or more of the following tests: FBG, 50 g GCT or 75 g/100 g OGTT, primarily between 24 and 28 weeks' gestation. There was large variation in study-defined thresholds or categories for elevated maternal glucose. Some studies based this on a specific glucose cutoff value, whereas others employed the use of pre-existing criteria, such as Carpenter and Coustan (CC) criteria or International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria and evaluated the number of abnormal values. Eleven studies reported on 2 groups: 1 with elevated and 1 with normal glucose tolerance (NGT), 5 studies included 3 groups: 2 elevated glucose groups compared to NGT, and 1 study reported on 4 groups based on glucose tolerance. All studies measured elevated glucose by discreet thresholds, rather than incremental increases in glucose (unlike the Farrar 2016 MA). "Elevated glucose" in the majority of studies cannot be compared with the definition of "elevated glucose" according to NICE criteria as the majority of included studies measured elevated glucose via the 50 g GCT followed by the 100 g OGTT, whereas NICE recommends a 2-hour 75 g OGTT. Therefore, the thresholds are not directly comparable (with the exception of FPG levels, as this is measured before a glucose dose is given). These studies were nevertheless

included because they initially included populations that would not be considered to be at risk of GDM based on specific risk factors. Only 2 studies, the MAMMA study and the GDMFU study could potentially be used to identify a new threshold above which there is increased risk of an outcome, because these were the only studies that used tests comparable to NICE and investigated single glucose thresholds (Table 6) in women without risk factors. In 9 additional studies that did not use the 75 g OGTT, the FPG cut-off values could be comparable to the NICE criteria. However, these cannot be used to identify a single threshold because they were part of criteria where elevated glucose was defined based on several possible tests (e.g. FPG, 50 g and/or 100g OGTT 1h, 2h and 3h values), and 1 or more abnormal value may have been required. In other words, not all women meeting the criteria for elevated glucose may have met the abnormal FPG threshold but still fulfilled the criteria based on other tests.

Full details of the thresholds and criteria are presented in Table 6. The majority of studies reported on both pregnancy and neonatal outcomes. Long-term outcomes were only reported in Farrar 2016 and in the supplementary publication on the HAPO Belfast cohort.

6. Summary of included studies for question 1

Study		Location	Design	Pregnant population	N in analysis	Glucose test(s)	Threshold(s) for elevated/ abnormal glucose	Timepoint of glucose measurement	Pregnancy outcomes	Neonatal outcomes	Longterm outcomes
Farrar 2016 HTA											
Farrar 2016 HTA and Born in Bradford (BiB) Study IPD[4, 55, 56]		Various (including nonOECD/ EEA)	SLR and MA	Pregnant women who had undergone assessment of glucose tolerance	NA	50 g GCT → 75 g or 100 g OGTT	**OR per 1 mmol/L increment**	Various (majority 24 to 28 weeks)	Y	Y	Y
Studies supplementary to Farrar 2016	**ATLANTIC-DIP** (Dennedy 2012)[53]	Ireland	Prospective cohort	Euthyroid women with singleton pregnancies	413	75 g OGTT	**OR per 1 mmol/L increment**	First trimester	Y	Y	N
	HAPO Belfast (Thaware 2015)[54]	Ireland	Prospective cohort	Offspring aged 5–7 years	1320	75 g OGTT	**OR per one unit rise in fasting, 1 h and 2 h OGTT**	28 weeks	N	N	Y
Studies identified from database searches											
Comparable with NICE criteria (2h 75 g OGTT [NICE thresholds: FPG ≥5.6 mmol/L or 2h OGTT ≥7.8 mmol/L])											
GDMFU (López del Val 2019)[57]		Spain	Retrospective cohort	Untreated mild GDM and nonGDM	1348	FPG	1) **5.1 mmol/L (92 mg/dL)**	24 to 28 weeks	Y	Y	N
MAMMA (Berntorp 2015)[58]		Sweden	Prospective cohort	Pregnant women representing different glucose categories	11,016	75 g OGTT	1) **2h 5.7 to 6.4 mmol/L** 2) **2h 6.5 to 7.2 mmol/L**	28 weeks	Y	Y	N
Include same test as NICE criteria or FPG but not comparable due to possibility of multiple different abnormal values											

Biri 2009[59]	Turkey	Retrospective cohort	All singleton pregnancies screened for GDM	2029	50 g GCT → 100 g OGTT	**1) Abnormal 50 g, normal 100 g:** 50 g abnormal 7.8 mmol/L (140 mg/dL) **2) One abnormal 100 g:** 100 g abnormal **FPG 5.8 mmol/L (105 mg/dL)** 1h 10.6 mmol/L (190 mg/dL) 2h 9.2 mmol/L (165 mg/dL) 3h 8.1 mmol/L (145 mg/dL)	24 to 28 weeks	Y	Y	N

Screening for Gestational Diabetes

Cheng 2009[60]	US	Retrospective cohort	All singleton pregnancies	14,693	50 g GCT → 100 g OGTT	**1) GDM by CC only (100 g):** **FPG 5.3 mmol/L (95 mg/dL)** **OR** 1h 10.0 mmol/L (180 mg/dL) 2h 8.6 mmol/L (155 mg/dL)	24 to 28 weeks	Y	Y	N

Study	Location	Design	Pregnant population	N in analysis	Glucose test(s)	Threshold(s) for elevated/ abnormal glucose	Timepoint of glucose measurement	Pregnancy outcomes	Neonatal outcomes	Longterm outcomes
						3h 7.8 mmol/L (140 mg/dL)				
Corrado 2009[61]	Italy	Retrospective cohort	Caucasian singleton pregnancies with positive screening test and OGTT	776	50 g GCT → 100 g OGTT	**1) GDM by CC only (100 g):** **FPG 5.3 mmol/L (95 mg/dL)** 1h 10.0 mmol/L (180 mg/dL) 2h 8.6 mmol/L (155 mg/dL) 3h 7.8 mmol/L (140 mg/dL)	24 to 28 weeks	Y	Y	N
Delibas 2018[62]	Turkey	Retrospective cohort	Singleton pregnancies with abnormal 1 h 50 g GCT	413	50 g GCT → 100 g OGTT	**1) Single high glucose value by NDDG criteria:** **FPG 5.3 mmol/L (95 mg/dL)** 1h 10.0 mmol/L (180 mg/dL) 2h 8.6 mmol/L (155 mg/dL) 3h 7.8 mmol/L (140 mg/dL)	24 to 28 weeks	Y	Y	N
Donovan 2017[63]	Canada	Retrospective cohort	All pregnancies	178,527	50 g GCT → 75 g OGTT	**1) HAPO 1.75:** **FPG ≥5.1 to <5.3 mmol/L** 1h ≥10 to <10.6 mmol/L **2h ≥8.5 to <9.0 mmol/L**	24 to 28 weeks	Y	Y	N

Jiang 2017[64]	Australia	Retrospective cohort	Singleton pregnancies with antenatal OGTT	4081	50 g GCT → 75 g OGTT	**1) GDM/IADPSG 2010Only:** FPG 5.1 to 5.4 mmol/L 2h 8.0 mmol/L	24 to 28 weeks	Y	Y	N

Screening for Gestational Diabetes

MFMU Network (Berggren 2012)[65]	US	Secondary analysis of RCT	Pregnancies with 1 h glucose load test result	1535	50 g GCT → 100 g OGTT	**1) Glucose intolerance (abnormal 50 g; normal 100g):** 1h 50g ≥7.5 to <11.1 mmol/L (≥135 to <200 mg/dL) **2) Mild untreated GDM (untreated):** ≥2 values above CC thresholds: FPG 5.3 mmol/L (95 mg/dL) 1h 10.0 mmol/L (180 mg/dL) 2h 8.6 mmol/L (155 mg/dL) 3h 7.8 mmol/L (140 mg/dL)	24 to 30 weeks	Y	Y	N
Meek 2015[66]	UK (England)	Retrospective cohort	All pregnancies	25,543	50 g GCT → 75 g OGTT	**1) GDM/IADPSG 2010only (NICE 2015negative):** FPG 5.1 to 5.5 mmol/L 1h ≥10.0 mmol/L 2h <7.8 mmol/L	26 to 28 weeks	Y	Y	N
Miyakoshi 2010[67]	Japan	Retrospective cohort	Singleton pregnancies	283	50 g GCT → 75 g OGTT	**1) 2 h IGT:** 8.3 mmol/L **2) 1 h IGT:** 10.0 mmol/L	24 to 27 weeks	Y	Y	N

Study	Location	Design	Pregnant population	N in analysis	Glucose test(s)	Threshold(s) for elevated/ abnormal glucose	Timepoint of glucose measurement	Pregnancy outcomes	Neonatal outcomes	Longterm outcomes
Not comparable with NICE criteria (different test or no FPG)										
Beksac 2018[68]	Turkey	Retrospective cohort	Singleton pregnancies	584	50 g GCT	1) 7.770 to <8.880 mmol/L; 2) 8.880 to 9.990 mmol/L; 3) >9.990 (n=20)	24 to 28 weeks	Y	Y	N
Berggren 2011[69]	US	Retrospective cohort	Women eligible for GDM screening	4659	50 g GCT → 100 g OGTT	**1) GDM by CC only:** 3h 7.8 mmol/L (140 mg/dL)	24 to 28 weeks	Y	Y	N

Davis 2018[70]	US	Retrospective cohort	Singleton pregnancies with glucose assessment	5973	50 g GCT → 100 g OGTT	**1) Mild hyperglycaemia:** Thresholds unclear **2) GDM/IADPSG 2010Only:** Thresholds unclear – abnormal 100 g OGTT based on IADPSG criteria (normal based on CC criteria)	24 to 28 weeks	Y	Y	N
Ezell 2015[71]	US	Prospective cohort	Black women aged 18 to 44	158	50 g GCT	1) 7.5 mmol/L (135 m/dL)	28 weeks	N	Y	N

Screening for Gestational Diabetes

LIFECODES (Noor 2019)[72]	US	Prospective cohort	Population with data on urinary phthalate metabolite concentrations and infants born ≥37 weeks' gestation	277	50 g GCT	1) 6.7 to < 7.8 mmol/L (120 to <140 mg/dL) 2) ≥7.8 mmol/L (≥140 mg/dL) without GDM	Second trimester	N	Y	N
Verd 2016[73]	Spain	Prospective cohort	Mother-infant dyads where mothers attempted breastfeeding	768	50 g GCT → 100 g OGTT	**1) MIGT:** 7.8 to <10.6 mmol/L	24 to 28 weeks	N	Y	N

Abbreviations: BiB, Born in Bradford Study; CC, Carpenter and Coustan criteria; EEA, European Economic Area; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; GDMFU; GDM Treatment Trial Follow-Up; HAPO, Hyperglycaemia and Adverse Pregnancy Outcomes; HTA, Health Technology Assessment; IADPSG; International Association of Diabetes and Pregnancy Study Groups; IGT, impaired glucose tolerance; IPD, individual patient data; LGA, large for gestational age; MA, meta-analysis; MFMU, Maternal-Fetal Medicine Units; MIGT, mild impairment of glucose tolerance; NA, not applicable; NDDG, National Diabetes Data Group; NR, not reported; OECD, Organisation for Economic Cooperation and Development; OGTT, oral glucose tolerance test; OR, odds ratio; RCT, randomised controlled trial; SLR, systematic literature review

Green font indicates a test used by NICE (either FPG or 2h 75 g OGTT) but a lower threshold than the NICE threshold. These tests identify a lower risk population compared with the current NICE screening tests. Red font indicates a test used by NICE but with the same or a higher threshold than the NICE threshold. Black font indicates a test is not used by NICE (for example, 50 g GCT or 100 g OGTT). Please note that some studies (Cheng 2009;[60] Corrado 2009;[61] Delibas 2018;[62] Donovan 2017;[63] Jiang 2017;[64] MFMU Network;[65] Meek 2015[66]) include green font but cannot be confirmed as including a low risk population (that is different to NICE) because women were included if they had abnormal value(s) on any of several different tests.

Discussion of findings

Quality assessments

Farrar 2016 SLR and MA

The quality of Farrar 2016 was appraised using the AMSTAR 2 checklist. Overall, the study quality was high, including clear objectives and eligibility criteria, a comprehensive search strategy and robust methodology (dual review). The results, including those from the MA, were clearly reported including a detailed discussion of the characteristics of included studies. The quality of the included studies was assessed using validated tools (Critical Appraisal Skills Programme [CASP] and Quality in Prognosis Studies [QUIPS]). Appropriate methods of statistical combination were used in the MA and risk of bias was accounted for in regression analyses while possible heterogeneity was examined using random-effects analyses. It was noted that there was considerable heterogeneity across studies assessing risk of macrosomia and LGA, however there was no evidence that the trend in risk associated with glucose level was different depending on the different glucose tests used.

It should be noted that the eligibility criteria for Farrar 2016 differed from this rapid review in that studies from any country were eligible, rather than being limited to OECD or EEA countries. Seven such studies were included in the MA, which may limit the generalisability of the results to a UK setting, although the vast majority of studies were from eligible countries.

Full details of the quality appraisal of Farrar 2016 are presented in Table 100 (Appendix 5).

Studies included in Farrar 2016

Farrar 2016 summarised that most studies were generally judged to be at low risk of bias. Selection of patients was not limited, there was little loss-to-follow-up and levels of glucose and outcomes were measured using standard criteria or definitions. The main potential risk of bias was a lack of blinding of participants and outcomes assessors to glucose levels, which may have resulted in outcome bias, in that assessors would have been aware of increased glucose levels and incorrectly attribute an outcome to this. There is an additional potential bias from confounding factors as studies did not adjust for maternal characteristics, such as maternal age or BMI, that may impact the risk of adverse neonatal

and maternal outcomes independently of glucose level. Most populations were from highincome countries, and would therefore be applicable to a UK setting.

Other included studies from database searches

The quality of the 17 primary studies identified in the database searches was appraised using the ROBINS-I tool; a summary is presented in Table 7 and the full appraisal is presented in Table 47 (Appendix 4). Overall, this evidence was at high risk of bias for confounding and at moderate risk of bias for outcome measurement and reporting. There was little concern or low risk of the other domains.

Table 7. Quality assessment of included studies

Study	Bias due to:							Overall risk of bias
	Confounding	Participant selection	Classification of interventions	Deviations from intended interventions	Missing data	Measurement of outcomes	Selection of the reported result	
Beksac 2018[68]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
Berggren 2011[69]	Moderate	Low	Low	Low	Low	Low	Moderate	Moderate
Berggren 2012 (MFMU)[65]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
Berntorp 2015 (Mamma study)[58]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
Biri 2009[59]	Serious	Low	Low	Low	Low	Low	Moderate	Serious
Cheng 2009[60]	Serious	Low	Low	Low	Low	Low	Moderate	Serious
Corrado 2009[61]	Serious	Low	Low	Low	Low	Low	Moderate	Serious
Delibas 2018[62]	Serious	Low	Low	Low	Low	Moderate	Moderate	Serious
Davis 2018[70]	Moderate	Low	Low	Low	Moderate	Moderate	Moderate	Moderate
Donovan 2017[63]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
Ezell 2015[71]	Serious	Low	Low	Low	Low	Low	Moderate	Serious
Jiang 2017[64]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
López del Val 2019 (GDMFU)[57]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
Meek 2015[66]	Serious	Low	Low	Low	Low	Low	Moderate	Serious
Miyakoshi 2010[67]	Moderate	Low	Low	Low	Low	Low	Moderate	Moderate
Noor 2019[72]	Moderate	Serious	Low	Low	Low	Low	Moderate	Serious
Verd 2016[73]	Serious	Serious	Low	Low	Low	Low	Moderate	Serious

Confounding

All studies were judged to be either at moderate or serious risk of bias in this domain. In the majority of studies, whilst women were not generally selected for specific risk factors, it was noted that the outcomes may have been influenced by uncontrolled maternal factors, such as age, ethnicity or BMI, the impacts of which on study quality were judged on a studybystudy basis. The majority of studies did adjust for such factors in statistical analyses, however 5 studies did not.[59, 61, 62, 72, 73] In 2 studies, women were specifically selected for the presence of risk factors. In the Maternal and Fetal Medicine Units (MFMU) Network study (Berggren 2012), the analysis set only included women who self-reported as either Hispanic or non-Hispanic White, and in Ezell 2015 only Black women were included.[65, 71] It should be noted that these studies were still included as Hispanic and Black are not specified as at-risk ethnicities for GDM.

Participant selection

All but 2 studies were judged to be at low risk of bias for participant selection. In the majority of cases, inclusion of participants was not based on outcomes or characteristics measured after hyperglycaemic status had been determined. However, selection bias may have been present in the LIFECODES study (Noor 2019), which only included women with available urinary phthalate metabolite concentration data, and Verd 2016, which only included women achieving term delivery who attempted breastfeeding.[72, 73] In most studies, all women received the glucose assessment at the same specified time period, usually in the region of 24 to 28 weeks' gestation, so this is not likely to have influenced selection bias.

Classification of interventions

All studies were judged to be at a low risk of bias in the classification of interventions domain. This was largely on the basis that the criteria for women classified as having elevated glucose were clearly defined (by specific thresholds or on the basis of pre-existing criteria) and the glucose assessment was performed before any outcome data were collected, so knowledge of outcomes could not have influenced the classification of women into glucose categories.

Deviations from intended interventions

Similarly, all studies were judged to be at low risk of bias in the deviations from intended interventions domain. In all studies, women in the eligible glucose categories were known not to have received any specific treatment for GDM.

Missing data

The majority of studies were judged to be at a low risk of bias due to missing data. The 1 study judged to be at moderate risk of bias excluded 755 participants due to lack of usable glucose data (Davis 2018), leaving 5973 women included in the analysis.[70] The remaining studies did not appear to exclude women on this basis or excluded only a small proportion of women, which was not expected to have affected the results, and were therefore judged to be at a low risk of bias.

Measurement of outcomes

Ten studies were judged to be at low risk of bias for the measurement of outcomes and 7 were judged to be at a moderate risk of bias. In all studies, it was judged that methods of outcome assessment were comparable across women in different glucose categories. For all studies, it was thought likely that outcomes assessors, for example midwives and obstetricians, would have been aware of a woman's glycaemic status, even though this was not explicitly reported in any study. Studies were judged to be at a moderate risk of bias when it was likely that systematic errors in the measurement of the outcome could have been introduced due to the outcomes assessor's awareness of the presence or absence of hyperglycaemia. Outcomes for which systematic errors were thought unlikely were unplanned C-section, measurement of LGA and macrosomia. For some studies it was also judged that systematic errors would not be introduced because the elevated glucose subgroup would have been considered as 'normal' in clinical practice at the time and thus assessors would not have been likely to perceive the women as having elevated glucose.

Selection of the reported result

All studies were judged to be at moderate risk of bias for selection of the reported result, as whilst preferential reporting of some outcomes or specific measures of an outcome was unlikely, this was unclear from what was reported in all studies. Furthermore, it was unclear if adjustments for confounding variables were pre-specified or selected based on the outcome results. Finally, for no study was there an *a priori* protocol or statistical analysis plan available.

Results

A study-level summary of data extracted from each included publication is presented in Appendix 3. Key results for all outcomes are presented in Table 15. For the purpose of

reporting results, pregnancy outcomes are grouped into (1) gestational age and pre-term birth and (2) pre-eclampsia and hypertension; neonatal outcomes are grouped into (1) stillbirth and perinatal mortality; (2) C-section, induction of labour and birth injury; (3) birth weight, macrosomia and LGA and (4) respiratory distress, congenital malformation, neonatal hypoglycaemia and admission to neonatal intensive care unit (NICU). Due to limited reporting, long-term outcomes are not grouped.

Pregnancy outcomes

The Farrar 2016 MA and the 15 included primary studies reported on at least 1 pregnancyassociated outcome.

Gestational age and pre-term birth

Gestational age at birth was reported by 10 studies (Table 8). Overall, there was no clear difference between elevated glucose and NGT. Values were consistently similar (in the region of 38 to 39 weeks), regardless of glucose level group. In the 7 studies reporting mean ± standard deviation (SD), this ranged from 38.5 ± 1.7 weeks[62] to 39.3 ± 2.0 weeks[70] in the NGT groups and from 38.4 ± 1.4 weeks[59] to 39.4 ± 1.9 weeks[70] in the elevated glucose groups. Four studies reported on levels of statistical significance. Of these, 2 found a significant difference between NGT and elevated glucose groups. Beksac 2018 found that median pregnancy duration was significantly longer for women with higher glucose according to the 50 g GCT (<7.77 mmol/L: 37 [30 to 41] weeks; 7.77 to <8.88 mmol/L: 37 [34 to 41] weeks, p=0.019; 8.88 to 9.99 mmol/L: 38.0 [31 to 40], p<0.001).[68] On the other hand, the MAMMA study reported that a higher proportion of women in the lowest glucose group, <5.7 mmol/L on 75 g OGTT, reached a gestational age of ≥42+0 weeks than women in higher glucose groups, 5.7 to 6.4 and 6.5 to 7.2 mmol/L (26.5% vs 24.7% vs 24.8%, p=0.006).[58] The 2 other studies that measured significance both found no significant difference in the NGT group compared to women with a single abnormal 100 g OGTT value.[61, 62]

Pre-term birth was reported by Farrar 2016 and 9 additional studies (Table 8), but results were inconsistent as to the risk of the outcome between elevated glucose and NGT groups. The MA in Farrar 2016 found that the OR of pre-term birth was 1.06 (95% CI 0.96 to 1.17), 0.77 (95% CI 0.62 to 0.96) and 1.07 (95% CI 0.99 to 1.15) per 1 mmol/L increment in glucose as measured by a 1 h 50 g GCT, FPG or 2 h 75 g/100 g OGTT respectively. For the most widely used test, OGTT, there was a trend towards a positive association between elevated glucose and pre-term birth, although this was not statistically significant. In the other studies, the proportion of pre-term births varied widely, from 0.4% to 20% across NGT groups and 0.6% to 26.2% in groups with abnormal glucose values.[58, 60] This

substantial variation is likely due to considerable heterogeneity between studies. Of the 4 studies that reported a p value for comparisons between NGT and elevated glucose groups, only the MAMMA study found that pre-term birth was significantly higher in women with elevated glucose (5.7 to 6.4 mmol/L or 6.5 to 7.2 mmol/L) compared with NGT (<5.7 mmol/L) (25.3% or 26.2% vs 20.0%; p=0.006).[58] This is noteworthy because the "elevated" glucose groups in the MAMMA study are still below the threshold considered by NICE to be abnormal (7.8 mmol/L). In those studies that reported ORs for between-group comparisons, different studies reported different directions of results. For example, whilst Davis 2018 reported 42% higher odds (OR 1.423, 95% CI 0.75 to 2.71),[70] Jiang 2017 reported 25% lower odds (OR 0.75, 95% CI 0.27 to 2.07) for the elevated glucose group compared with NGT, however, this may in part be owing to a relatively small sample size.[64] Coupled with inconsistent ORs from Farrar 2016 and wide CIs in all cases, there was no clear direction of effect in any study for elevated glucose on pre-term birth (Table 8).

In summary, there was no clear, consistent association between gestational age or preterm birth and elevated glucose. While the MAMMA study did compare different 75 g OGTT glucose categories below the threshold considered by NICE, demonstrating a higher risk of pre-term birth from 5.7 mmol/L upwards, no other studies allowed for the identification of a clear glucose threshold where risk of pre-term birth or decreased/increased gestational age may be differentiated.

Table 8. Gestational age and pre-term birth

Outcome	Study	Glucose test	Glucose threshold	Outcome unit	Outcome value	Risk (95% CI)	pvalue
Gestational age at birth	Beksac 2018	50 g GCT	<7.770 mmol/L (n=352) 7.770 to <8.880 mmol/L (n=165) 8.880 to 9.990 mmol/L (n=47) >9.990 (n=20)	Median weeks (range)	37.0 (30 to 41) 37.0 (34 to 41) 38.0 (31 to 40) 37.5 (36 to 40)	NR	Ref **0.019** **<0.001** NS
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767) Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)	Mean weeks (SD)	Hispanic: 39.4 (1.6) Non-Hispanic white: 39 (1.5) Hispanic: 39.2 (1.6) Non-Hispanic white: 38.7 (1.9)	NR	NR
	MAMMA (Berntorp 2015)	75 g OGTT	<5.7 mmol/L (n=2637) 5.7 to 6.4 mmol/L (n=2783) 6.5 to 7.2 mmol/L (n=2819)	n (%)	**37–41+6 weeks: 2345 (24.0)** **≥42+0 weeks: 175 (26.5)** **37–41+6 weeks: 2472 (25.3)** **≥42+0 weeks: 163 (24.7)** **37–41+6 weeks: 2502 (25.6)** **≥42+0 weeks: 164 (24.8)**	NR	**0.006**
	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941) Mild hyperglycaemia (threshold unclear) (n=544) GDM/IADPSG 2010-Only (thresholds unclear) (n=181)	Mean weeks (SD)	39.3 (2.0) 39.3 (2.0) 39.4 (1.9)	NR	NR
	Delibas 2018	50 g GCT → 100 g OGTT	NGT (n=316) Single high glucose value: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=33)	Mean weeks (SD)	**38.5 (1.7)** 38.5 (1.3)	NR	NS
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406) GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)	Mean weeks (95% CI)	39.3 (39.3 to 39.4) 39.1 (38.9 to 39.2)	NR	NR
	Biri 2009	50 g GCT → 100 g OGTT	NGT (n=1432) Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326) One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)	Mean weeks (SD)	39.0 (1.4) 38.6 (1.3) 38.4 (1.4)	NR	NR

	Corrado 2009	50 g GCT → 100 g OGTT	One abnormal 100 g: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=152)	Mean weeks (SD)	38.5 (1.8)	NR	0.06
			NGT (n=624)		38.8 (1.6)		
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	Median weeks (range)	39.3 (38.1 to 40.3)	NR	NR
			NGT (n=3117)		39.3 (38.1 to 40.4)		
	Miyakoshi 2010	50 g GCT →	NGT (n=4512)		38.7 (1.9)	NR	NR
		75 g OGTT	2 h IGT (8.3 mmol/L) (n=108)	Mean weeks (SD)	38.5 (2.1)		
			1 h IGT (10.0 mmol/L) (n=66)		38.6 (1.6)		
Pre-term birth	Farrar 2016 HTA	50 g GCT	Per 1 mmol/L increment	NA	NA	OR 1.06 (0.96 to 1.17)	NR
		75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	NA	FPG: **OR 0.77 (0.62 to 0.96)** 1 h: NR 2 h: OR 1.07 (0.99 to 1.15)	NR

Outcome	Study	Glucose test	Glucose threshold	Outcome unit	Outcome value	Risk (95% CI)	pvalue
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767)	n (%)	Hispanic: 35 (7) Non-Hispanic white: 14 (6)	NR	NR
			Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)		Hispanic: 23 (9) Non-Hispanic white: 134 (12)		
	MAMMA (Berntorp 2015)	75 g OGTT	<5.7 mmol/L (n=2637)	n (%)	**117 (20.0)**	NR	**0.006**
			5.7 to 6.4 mmol/L (n=2783)		**148 (25.3)**		
			6.5 to 7.2 mmol/L (n=2819)		**153 (26.2)**		
	Jiang 2017	50 g GCT → 75 g OGTT	NGT (n=3185)	n (%)	178 (5.6)	OR 0.75 (0.27 to 2.07)	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.4 mmol/L, 2h 8.0 mmol/L (n=94)		4 (4.3)		
	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941)	n (%)	455 (9.2)	OR 1 (ref)	Ref
			Mild hyperglycaemia (threshold unclear) (n=544)		51 (9.4)	OR 1.020 (0.75 to 1.38) AOR 1.243 (0.83 to 1.86)	0.899 0.289 0.673 0.284

					15 (8.3)	OR 0.891 (0.52 to 1.52) AOR 1.423 (0.75 to 2.71)	
			GDM/IADPSG 2010-Only (thresholds unclear) (n=181)				
	Delibas 2018	50 g GCT → 100 g OGTT	NGT (n=316)	n (%)	4 (1.4)	NR	NS
			Single high glucose value: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=33)		1 (3.3)		
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406)	n (%)	127 (5.3)	NR	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)		29 (7.5)		
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	n (%)	NR (0.4)	AOR 1.36 (0.84 to 2.18)	0.09
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)		NR (0.6)		
	Biri 2009	50 g GCT →	NGT (n=1432)	n (%)	NR (1.4)	NR	NR
		100 g OGTT	Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326)		NR (9.5)		
			One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)		NR (7.0)		
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	n (%)	66 (14)	APR 1.09 (0.86 to 1.39)	NR
			NGT (n=3117)		403 (13)		

Bolded results are indicated as statistically significant at p<0.05.

Abbreviations: AOR, adjusted odds ratio; APR, adjusted prevalence ratio; CC, Carpenter and Coustan; CI, confidence interval; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; HTA, Health Technology Assessment; IADPSG, International Association of Diabetes and Pregnancy Study Groups; IGT, impaired glucose tolerance; MFMU, Maternal-Fetal Medicines Unit; NGT, normal glucose tolerance; NA, not applicable; NR, not reported; NS, not significant; OGTT, oral glucose tolerance test; OR, odds ratio; SD, standard deviation.

Pre-eclampsia and hypertension

Eight studies, including the Farrar 2016 SLR, reported on pre-eclampsia, 6 on hypertension and 2 on pre-eclampsia or hypertension (Table 9).

Rates of pre-eclampsia were generally low in all studies and similar across all glucose groups (0% to 7.2% for NGT and 0% to 13% for elevated glucose). Farrar 2016 reported that for each 1 mmol/L increment of glucose the OR of pre-eclampsia was increased; this ranged from 1.19 (95% CI 1.15 to 1.24) for glucose measured at 1 h (either 75 g or 100 g OGTT) to 2.15 (95% CI 1.45 to 3.19) for fasting glucose in the same test.[4] Only 1 other study found pre-eclampsia more likely in the elevated glucose group. Berggren 2011 reported an adjusted prevalence ratio (APR) of 1.47 (95% CI 1.02 to 2.13) for the elevated glucose group (defined as untreated GDM according to the CC criteria) compared with NGT.[69] By contrast, Cheng 2009 reported a statistically non-significant adjusted OR (AOR) of 1.30 (95% CI 0.71 to 2.38) for the elevated glucose group (similarly defined as untreated GDM according to the CC criteria) compared to NGT.[60] The GDMFU study also reported an OR of 1.02 (95% CI 0.28 to 3.75) for women with glucose levels of ≥5.1 mmol/L (classified as elevated) compared to <5.1 mmol/L (classified as NGT).[57] The 5.1 mmol/L value was based on FPG, rather than the more commonly employed 75 g or 100 g glucose bolus dose, so is not directly comparable to the OGTT tests.

Results for pregnancy-induced hypertension were also varied. Where reported, ORs ranged from 1.053 (95% CI 0.78 to 1.43; p=0.740)[70] to 1.5 (95% CI 1.4 to 1.7; p<0.01)[63] for elevated glucose compared to NGT. In the supplementary ATLANTIC-DIP publication (included here as this outcome was not reported separately by Farrar 2016), the ORs per 1 mmol/L increment of glucose were also inconclusive. ORs were 1.220 (95% CI 0.663 to 2.246), 1.049 (95% CI 0.910 to 1.209) and 1.160 (95% CI 0.960 to 1.402) for FPG, 1 h 75 g OGTT and 2 h 75 g OGTT, respectively.[53]

In summary, of 8 studies reporting on pre-eclampsia, a statistically higher risk of the outcome among women with abnormal glucose tolerance was only shown by 2 studies, the others either not reporting a statistical comparison or reporting it to be statistically nonsignificant (including the GDMFU trial, which could have identified a potential threshold for elevated risk). It is noteworthy that 1 of the significant results is from the MA by Farrar. Out of 6 studies reporting hypertension in pregnancy, 3 found that groups with abnormal glucose were more at risk whereas the other 3 found no significant increase in risk or did

not report a statistical comparison. Given the above, it is unclear if there is an increased risk of pre-eclampsia or hypertension in women with decreased glucose tolerance and it is not possible to identify a specific threshold at which risk is increased.

Table 9. Pre-eclampsia and hypertension

Outcome	Study	Glucose test	Glucose threshold	Outcome value (n [%])	Risk (95% CI)	p-value
Pre-eclampsia	Farrar 2016 HTA	50 g GCT	Per 1 mmol/L increment	NA	**OR 1.25 (1.13 to 1.39)**	NR
		75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	FPG: **OR 2.15 (1.45–3.19)** 1 h: **OR 1.19 (1.15–1.24)** 2 h: **OR 1.23 (1.18– 1.29)**	NR
	GDMFU (López del Val 2019)	FPG	<5.1 mmol/L (n=1193)	19 (1.7)	OR 1.02 (0.28 to 3.75)	NS
			≥5.1 mmol/L (n=155)	3 (2.1)		
	Delibas 2018	50 g GCT → 100 g OGTT	NGT (n=316)	4 (1.4)	NR	NR
			Single high glucose value: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=33)	1 (3.3)		
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406)	174 (7.2)	NR	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)	39 (10.1)		
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	NR (4.5)	AOR 1.30 (0.71 to 2.38)	NR
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)	NR (6.2)		
	Biri 2009	50 g GCT → 100 g OGTT	NGT (n=1432)	NR (1.5)	NR	NR
			Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326)	NR (2.3)		
			One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)	NR (2.1)		
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	53 (13)	APR **1.47 (1.02 to 2.13)**	Significant
			NGT (n=3117)	264 (8)		
	Miyakoshi 2010	50 g GCT → 75 g OGTT	NGT (n=4512)	NR (1.8)	NR	NR
			2 h IGT (8.3 mmol/L) (n=108)	NR (0.9)		
			1 h IGT (10.0 mmol/L) (n=66)	NR (0)		
Pregnancy-induced hypertension	ATLANTIC-DIP (Dennedy 2012)	75 g OGTT	Per 1 mmol/L increment (FPG)	NA	OR 1.220 (0.663 to 2.246)	NR
			Per 1 mmol/L increment (1 h glucose)		OR 1.049 (0.910 to 1.209)	
			Per 1 mmol/L increment (2 h glucose)		OR 1.160 (0.960 to 1.402)	
	Donovan 2017	50 g GCT →	Normal 50 g screen (n=144,191)	8028 (5.6)	OR 1 (ref)	Ref

		75 g OGTT	Normal 75 g screen (n=21,248)	1550 (73)	OR **1.3 (1.2 to 1.4)**	**<0.01**
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)	390 (9.1)	OR **1.5 (1.4 to 1.7)**	**<0.01**
	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941)	442 (8.9)	OR 1 (ref)	Ref
			Mild hyperglycaemia (threshold unclear) (n=544)	51 (9.4)	OR 1.053 (0.78 to 1.43) AOR 1.080 (0.70 to 1.66)	0.740 0.723 0.141 0.563
			GDM/IADPSG 2010-Only (thresholds unclear) (n=181)	22 (12.2)	OR 1.409 (0.89 to 2.22) AOR 1.215 (0.63 to 2.35)	

Outcome	Study	Glucose test	Glucose threshold	Outcome value (n [%])	Risk (95% CI)	p-value
	Corrado 2009	50 g GCT → 100 g OGTT	One abnormal 100 g: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=152)	21(13.8)	NR	**0.001**
			NGT (n=624)	27 (4.3)		
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	33 (7)	APR **1.48 (1.02 to 2.13)**	NR
			NGT (n=3117)	150 (5)		
	Miyakoshi 2010	50 g GCT → 75 g OGTT	NGT (n=4512)	NR (1.9)	NR	NR
			2 h IGT (8.3 mmol/L) (n=108)	NR (2.8)		
			1 h IGT (10.0 mmol/L) (n=66)	NR (4.6)		
Pre-eclampsia or hypertension	Farrar 2016 HTA	50 g GCT	Per 1 mmol/L increment	NA	OR 1.02 (0.75 to 1.38)	NR
		75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	FPG: **OR 1.91 (1.49 to 2.43)** 1 h: NR 2 h: **OR 1.19 (1.08–1.30)**	NR
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767)	Hispanic: 38 (7) Non-Hispanic white: 27 (11)	NR	NR
			Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)	Hispanic: 37 (15) Non-Hispanic white: 13 (11)		

Bolded results are indicated as statistically significant at p<0.05.

Abbreviations: AOR, adjusted odds ratio; APR, adjusted prevalence ratio; CC, Carpenter and Coustan; CI, confidence interval; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; GDMFU, GDM Treatment Trial Follow-Up; HAPO, Hyperglycaemia and Adverse Pregnancy Outcomes; HTA, Health Technology Assessment; IADPSG, International Association of Diabetes and Pregnancy Study Groups; IGT, impaired glucose tolerance; MFMU, Maternal-Fetal Medicines Unit; NGT, normal glucose tolerance; NA, not applicable; NR, not reported; NS, not significant; OGTT, oral glucose tolerance test; OR, odds ratio; SD, standard deviation.

Neonatal outcomes

The Farrar 2016 MA and 18 studies identified from the database searches reported on at least 1 neonatal outcome.

Stillbirth and perinatal mortality

Four studies reported on stillbirth or perinatal mortality (n=3 and n=1, respectively) (Table 10).[63, 64, 66] For all studies, rates of the events were very low across all groups (0% to 0.3% for stillbirth; 0.1% for perinatal mortality). Furthermore, any potential differences cannot be quantified as no study reported on measures of risk or levels of statistical significance. The 1 study that reported on perinatal mortality, Donovan 2017, also reported on stillbirth and included large numbers of women in all glucose groups (21,248 to 144,191 women in the normal glucose groups and 4308 in the elevated glucose group) only finding 599 deaths, indicating that perinatal mortality and stillbirth are rare events and even the largest studies may be underpowered to detect a difference in these outcomes.[63]

Based on the identified evidence, a glucose threshold above which the risk of these outcomes would increase cannot be identified. The low number of events in both NGT and elevated glucose groups may indicate that there is no association between glucose and stillbirth or perinatal mortality.

Table 10. Stillbirth and perinatal mortality

Outcome	Study	Glucose test	Glucose threshold	Outcome value, n (%)	Risk (95% CI)	p-value
Stillbirth	Donovan 2017	50 g GCT → 75 g OGTT	Normal 50 g screen (n=144,191)	343 (0.2)	NR	NR
			Normal 75 g screen (n=21,248)	65 (0.3)		
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)	13 (0.3)		
	Jiang 2017	50 g GCT → 75 g OGTT	NGT (n=3185)	0 (0.3)	-	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.4 mmol/L, 2h 8.0 mmol/L (n=94)	0 (0)		
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406)	5 (0.2)	NR	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)	1 (0.3)		
Perinatal mortality	Donovan 2017	50 g GCT → 75 g OGTT	Normal 50 g screen (n=144,191)	150 (0.1)	NR	NR
			Normal 75 g screen (n=21,248)	22 (0.1)		
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)	6 (0.1)		

Abbreviations: CI, confidence interval; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; HAPO, Hyperglycaemia and Adverse Pregnancy Outcomes; IADPSG, International Association of Diabetes and Pregnancy Study Groups; NGT, normal glucose tolerance; NR, not reported; OGTT, oral glucose tolerance test.

C-section, induction of labour and birth injury

Neither GDMFU nor the MAMMA study reported on C-section or induction of labour, therefore no potential thresholds for elevated risk could be identified. Nevertheless, twelve studies including Farrar 2016 reported on rates of C-section, reporting mixed results on whether elevated glucose leads to a higher risk of this outcome. Rates ranged from 16.9% to 54.8% in the NGT groups, and 22.7% to 63.4% in elevated glucose groups.[59, 60] In addition to the Farrar 2016 MA, 1 other study reported ORs per 1 mmol/L increment in glucose (based on 1 h 50 g GCT).[71] Whilst Farrar 2016 reported an OR of 1.35 (95% CI 1.24 to 1.49) per 1 mmol/L increment, Ezell 2015 reported a range of AORs for overall, parous and nulliparous populations, all of which were either not statistically significant (p>0.05) or were under 1.05 (95% CI 1.00 to 1.05; AOR for nulliparous women, adjusted for maternal age, pre-pregnancy BMI and prior C-section).[71] The same study also reported AORs for the comparison between 50 g GCT <7.5 mmol/L (NGT) and ≥7.5 mmol/L (elevated glucose), also finding no association for nulliparous women and in AOR with large CIs for parous women (5.1, 95% CI 0.7 to 37.4; p=0.113). However, of the 5 other studies reporting measures of risk, all reported at least 1 comparison where the odds of C-section were significantly higher for women with elevated glucose than NGT (Table 11).[60, 63, 64, 69, 70]

Three studies reported on induction of labour, all of which presented results consistent with induction of labour being greater in women with elevated glucose. In the Farrar 2016 MA, ORs per 1 mmol/L increment of glucose were 1.31 (95% CI 1.14 to 1.50) and 1.10 (95% CI 1.04 to 1.16) for FPG and 2 h 75 g/100g OGTT, respectively.[4] In Donovan 2017, women in the HAPO 1.75 group (defined as at least 1 abnormal value on the 75 g OGTT corresponding to glucose values that were associated with an AOR of 1.75 for specified adverse events in the HAPO study) had an OR of 1.1 (95% CI 1.0 to 1.2; p<0.01) compared to women with a normal 50 g glucose screen.[63] Similarly, the frequency of induction of labour was significantly higher in women with GDM as defined by the CC criteria compared with those with NGT (32% vs 25%; p-value 'significant').[69]

Several studies reported subsets of outcomes related to birth injury, including trauma during vaginal delivery (n=2), shoulder dystocia (n=5) and 3rd or 4th degree lacerations (n=3). Trauma during vaginal delivery was reported as significantly lower for women with NGT (FPG<5.1 mmol/L) compared with elevated (FPG≥5.1 mmol/L) glucose in the GDMFU study (OR 3.10, 95% CI 1.15 to 8.32; p=0.02),[57] but this was not significantly different in another study (Cheng 2009) reporting this outcome (AOR 1.26, 95% CI 0.66 to 2.42; p=0.43).[60] The number of included women was higher in Cheng 2009 (13,940 women) compared with GDMFU (1193 women), and the OR in Cheng 2009 was adjusted for potential confounding

factors: parity, maternal age, race or ethnicity, gestational weight gain, gestational age at delivery, year of delivery, epidural anaesthesia and induction of labour; therefore, the estimate from Cheng 2009 may be more robust. This is particularly of note given that the GDMFU study reported on a potential threshold for elevated glucose, suggesting that it may not be reliable to draw conclusions on FPG 5.1 mmol/L being a threshold for increased risk, based on the GDMFU study alone.[57, 60]

For shoulder dystocia, the results from Farrar 2016 per 1 mmol/L increment of glucose supported an increased risk of outcome with increasing glucose (ORs ranging from 1.26, 95% CI 1.10 to 1.43 for 1 h 50 g GCT, to 1.97, 95% CI 1.36 to 2.85 for FPG). Similarly, Cheng 2009 reported an AOR of 2.24 (95% CI 1.03 to 4.88) for shoulder dystocia in women with elevated glucose.[60] However, similar results were not found in the other 3 studies reporting this outcome, with ORs ranging from 0.540 (95% CI 0.19 to 1.50; p=0.236) to 1.592 (95% CI 0.69 to 3.68; p=0.276).[64, 69, 70] Of note, the lower and upper range of ORs were reported in the same study in this case, which compared 2 categories of elevated glucose with NGT, suggesting that the way elevated glucose is defined may have a strong effect on the risk to shoulder dystocia.[70] No significant differences were reported in any of the 3 studies reporting on lacerations (p>0.05), with measures of risk ranging from an APR of 0.83 (95% CI 0.48 to 1.44) to an OR of 1.655 (95% CI 0.91 to 3.02; p=0.101) for elevated glucose compared to NGT (Table 11).

Other outcomes related to parturition that were reported included spontaneous vertex delivery, instrumental delivery and postpartum haemorrhage, showing no difference between elevated glucose and NGT groups. Full details of these results are presented in Appendix 2.

In summary, while there was an association between elevated glucose and increased risk of C-section and induction of labour, no specific glucose threshold risk was identified as this was not reported in any study looking at specific thresholds. Associations between glucose and risk were inconsistent for should dystocia, trauma during delivery and lacerations.

Table 11. C-section, induction of labour and birth injury

Outcome	Study	Glucose test	Glucose threshold	Outcome value, n (%)	Risk (95% CI)	p-value
C-section	Farrar 2016 HTA	50 g GCT	Per 1 mmol/L increment	NA	OR **1.35 (1.23 to 1.49)**	NR
		75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	FPG: OR **1.59 (1.49 to 1.70)** 1 h: OR **1.18 (1.15 to 1.20)** 2 h: OR 1.10 (0.96 to 1.25)	NR
	Donovan 2017	50 g GCT → 75 g OGTT	Normal 50 g screen (n=144,191)	37,455 (26.0)	OR 1 (ref)	Ref
			Normal 75 g screen (n=21,248)	6535 (30.8)	OR **1.2 (1.1 to 1.2)**	**<0.01**
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)	1561 (36.2)	OR **1.4 (1.3 to 1.5)**	**<0.01**
	Jiang 2017	50 g GCT → 75 g OGTT	NGT (n=3185)	536 (20.1)	OR **2.03 (1.23 to 3.35)**	**p<0.05**
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.4 mmol/L, 2h 8.0 mmol/L (n=94)	24 (33.8)		
	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941)	1267 (25.6)	OR 1 (ref)	Ref
			Mild hyperglycaemia (threshold unclear) (n=544)	175 (32.2)	OR **1.375 (1.14 to 1.66)** AOR 1.181 (0.91 to 1.52)	**0.001** 0.202
			GDM/IADPSG 2010-Only (thresholds unclear) (n=181)	51 (28.2)	OR 1.138 (0.82 to 1.58) AOR 0.810 (0.51 to 1.29)	0.444 0.377
	Delibas 2018	50 g GCT → 100 g OGTT	NGT (n=316)	90 (28.5)	NR	NR
			Single high glucose value: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=33)	11 (33.3)		
	Ezell 2015	50 g GCT	<7.5 mmol/L (n unclear)	NR	Parous: AOR 5.1 (0.7 to 37.4) Nulliparous: no association	0.113
			≥7.5 mmol/L (n unclear)	NR		NR
			Per 1 mmol/L increment (1 h glucose)	NA	Overall: OR 1.01 (1.00 to 1.03) Overall: AOR 1.01 (1.00 to 1.03) Parous: OR 1.00 (0.98 to 1.02) Parous: AOR 1.00 (0.98 to 1.02) Nulliparous: **OR 1.03 (1.00 to 1.05)** Nulliparous: **AOR 1.05 (1.00 to 1.05)**	0.131 0.356 0.856 0.884 **0.034** **0.029**
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406)	473 (19.7)	NR	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)	94 (24.3)		
	Verd 2016	50 g GCT →	NGT (n=616)	NR (79)	NR	0.67

		100 g OGTT	MIGT (7.8 mmol/L to <10.6 mmol/L) (n=152)	NR (21)	NR	
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	NR (16.9)	AOR **1.44 (1.01 to 2.07)**	**<0.001**
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)	NR (22.7)		
	Biri 2009	50 g GCT →	NGT (n=1432)	NR (54.8)	NR	NR

		100 g OGTT	Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326)	NR (63.1)		
			One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)	NR (63.4)		
	Corrado 2009	50 g GCT → 100 g OGTT	One abnormal 100 g: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L,	85 (56)	NR	**0.0001**

Outcome	Study	Glucose test	Glucose threshold	Outcome value, n (%)	Risk (95% CI)	p-value
			3h 7.8 mmol/L (n=152)			
			NGT (n=624)	243 (39)		
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	160 (35)	APR **1.16 (1.04 to 1.30)**	NR
			NGT (n=3117)	942 (30)		
Induction of labour	Farrar 2016 HTA	75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	FPG: OR **1.31 (1.14 to 1.50)** 1 h: NR 2 h: OR **1.10 (1.04 to 1.16)**	NR
	Donovan 2017	50 g GCT → 75 g OGTT	Normal 50 g screen (n=144,191)	39,611 (27.5)	OR 1 (ref)	Ref
			Normal 75 g screen (n=21,248)	5887 (27.7)	OR 1.0 (1.0 to 1.0)	0.47
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)	1274 (29.6)	OR **1.1 (1.0 to 1.2)**	**<0.01**
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	149 (32)	NR	**Significant**
			NGT (n=3117)	772 (25)		
Trauma during vaginal delivery	GDMFU (López del Val 2019)	FPG	<5.1 mmol/L (n=1193)	19 (1.6)	OR **3.10 (1.15 to 8.32)**	**0.02**
			≥5.1 mmol/L (n=155)	9 (5.7)		
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	NR (3.7)	AOR 1.26 (0.66 to 2.42)	0.43
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)	NR (4.4)		
Shoulder dystocia	Farrar 2016 HTA	50 g GCT	Per 1 mmol/L increment	NA	OR **1.26 (1.10 to 1.43)**	NR
		75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	FPG: OR **1.97 (1.36 to 2.85)** 1 h: NR 2 h: OR **1.38 (1.22 to 1.56)**	NR
	Jiang 2017		NGT (n=3185)	215 (6.8)	OR 0.78 (0.31 to 1.93)	NR

		50 g GCT → 75 g OGTT	GDM/IADPSG 2010-Only: FPG 5.1 to 5.4 mmol/L, 2h 8.0 mmol/L (n=94)	5 (5.3)		
	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941)	104 (2.1)	OR 1 (ref)	Ref
			Mild hyperglycaemia (threshold unclear) (n=544)	11 (2.0)	OR 0.953 (0.51 to 1.79) AOR 0.540 (0.19 to 1.50)	0.880 0.236 0.276
			GDM/IADPSG 2010-Only (thresholds unclear) (n=181)	6 (3.4)	OR 1.592 (0.69 to 3.68) AOR 1.294 (0.40 to 4.21)	0.669
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	NR (1.7)	AOR **2.24 (1.03 to 4.88)**	NR
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)	NR (3.3)		
	Berggren 2011	50 g GCT →	GDM by CC only (3h 7.8 mmol/L) (n=460)	24 (5)	APR 1.41 (0.91 to 2.18)	NR

		100 g OGTT	NGT (n=3117)	109 (4)		
Lacerations (3rd/4th degree)	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941)	203 (4.2)	OR 1 (ref)	Ref
			Mild hyperglycaemia (threshold unclear) (n=544)	30 (5.5)	OR 1.352 (0.91 to 2.01) AOR 1.024 (0.59 to 1.78)	0.134 0.934
			GDM/IADPSG 2010-Only (thresholds unclear) (n=181)	12 (6.7)	OR 1.655 (0.91 to 3.02) AOR 0.925 (0.33 to 2.58)	0.101 0.882
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	NR (9.0)	AOR 1.16 (0.73 to 1.86)	0.14
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)	NR (11.4)		
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	14 (3)	APR 0.83 (0.48 to 1.44)	NR
			NGT (n=3117)	118 (4)		

Bold results are significant at p<0.05

Abbreviations: AOR, adjusted odds ratio; APR, adjusted prevalence ratio; CC, Carpenter and Coustan; CI, confidence interval; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; GDMFU, GDM Treatment Trial Follow-Up; HAPO, Hyperglycaemia and Adverse Pregnancy Outcomes; HTA, Health Technology Assessment;

IADPSG, International Association of Diabetes and Pregnancy Study Groups; MIGT, mild impairment of glucose tolerance; NGT, normal glucose tolerance; NR, not reported; OGTT, oral glucose tolerance test.

Macrosomia, LGA and birth weight

Eleven studies including Farrar 2016 reported on macrosomia. The majority of those reporting a measure of risk found a strong association between elevated glucose and macrosomia with ORs ranging from 1.876 (95% CI 1.08 to 3.25; p=0.025)[70] to 4.47 (95% CI 2.26 to 8.86; p=0.01).[60] This included the GDMFU study, which found that a FPG threshold of ≥5.1 mmol/L was associated with a significantly higher rates of macrosomia than below this threshold when using an unadjusted odds ratio (p<0.05).[57] Similar results were reported by Farrar 2016 with ORs of 1.14 (95% CI 1.10 to 1.18) to 2.06 (95% CI 1.86 to 2.28) per 1 mmol/L increment as measured by the 1 h 50 g GCT and FPG, respectively.[4] However, no association was found between 1 of the elevated glucose groups in Davis 2018, classified as 'mild hyperglycaemia', and NGT (OR 1.196, 95% CI 0.90 to 1.59; p=0.222; AOR 0.988, 95% CI 0.66 to 1.48; p=0.955).[70] Moreover, when adjusted for BMI, age and previous GDM, the OR reported in the GDMFU study for the elevated glucose group decreased from 2.42 (95% CI 1.27 to 4.62; p<0.05) to 1.50 (95% CI 0.63 to 3.57; p=0.3).[57] Excluding one study from Japan, where rates of macrosomia were reported as 0% for elevated glucose and 0.7% for NGT,[67] proportions in the elevated glucose groups ranged from 4.0% to 28.9% compared with 1.6% to 16.8% in the NGT groups.[60, 66]

Twelve studies including Farrar 2016 reported on LGA. Similarly to the results for macrosomia, the majority of those reporting on measures of risk found significant associations between elevated glucose and LGA, ranging from OR 1.09 (95% CI 1.01 to 1.18; p=0.028)[58] to 4.28 (95% CI 2.24 to 8.18; p<0.001).[60] This included the MAMMA study, which found a significantly higher rate of LGA in women meeting a threshold of 6.5 to 7.2 mmol/L on a 75 g OGTT test compared with thresholds of >5.7 mmol/L or 5.7 to 6.4 mmol/L (p<0.001).[58] Cheng 2009 reported the highest ORs for both macrosomia and LGA (see Table 12).[60]

Studies reporting on birth weight (n=6) were less informative than those reporting on macrosomia or LGA, with no measures of mean difference presented in any study. One study found no significant difference between median birth weights in the NGT and elevated glucose groups (3272 g vs 3395 g; p=0.018)[73] whilst another found that mean birth weight was significantly lower for women with reactive hypoglycaemia compared with NGT and women with a single high glucose value (2852.0 ± 544.6 g vs 3282.4 ± 452.8 g or 3290.6 ± 510.5 g; p<0.05).[62]

Table 12. Macrosomia, LGA and birth weight

Outcome	Study	Glucose test	Glucose threshold	Outcome unit	Outcome value	Risk (95% CI)	p-value
Macrosomia	Farrar 2016 HTA	50 g GCT	Per 1 mmol/L increment	NA	NA	OR 1.14 (1.10 to 1.18)	NR
		75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	NA	FPG: OR 2.06 (1.86 to 2.28) 1 h: NR 2 h: OR 1.21 (1.16 to 1.26)	NR
	GDMFU (López del Val 2019)	FPG	<5.1 mmol/L (n=1193)	n (%)	40 (4.3)	OR 1 (ref)	Ref
			≥5.1 mmol/L (n=155)		12 (7.2)	OR 2.42 (1.27 to 4.62) AOR[a] 1.50 (0.63 to 3.57)	<0.05 0.3
	Donovan 2017	50 g GCT → 75 g OGTT	Normal 50 g screen (n=144,191)	n (%)	13,924 (9.5)	NR	NR
			Normal 75 g screen (n=21,248)		2385 (11.0)		
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)		594 (13.5)		
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767)	n (%)	Hispanic: 62 (12) Non-Hispanic white: 23 (9)	NR	NR
			Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)		Hispanic: 40 (16) Non-Hispanic white: 17 (15)		
	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941)	n (%)	455 (9.2)	OR 1 (ref)	Ref
			Mild hyperglycaemia (threshold unclear) (n=544)		59 (10.8)	OR 1.196 (0.90 to 1.59) AOR 0.988 (0.66 to 1.48)	0.222 0.955
			GDM/IADPSG 2010-Only (thresholds unclear) (n=181)		32 (17.8)	OR 2.126 (1.43 to 3.15) AOR 1.876 (1.08 to 3.25)	0.0002 0.025
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406)	n (%)	403 (16.8)	NR	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)		112 (28.9)		
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	n (%)	NR (1.6)	AOR 4.47 (2.26 to 8.86)	0.01
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)		NR (4.0)		
	Biri 2009	50 g GCT → 100 g OGTT	NGT (n=1432)	n (%)	NR (5.8)	NR	NR
			Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326)		NR (8.3)		
			One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)		NR (12.7)		
	Corrado 2009	50 g GCT → 100 g OGTT	One abnormal 100 g: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=152)	n (%)	19 (12.5)	NR	0.01
			NGT (n=624)		39 (6.2)		

	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	n (%)	78 (17)	APR 1.25 (1.01 to 1.56)	Significant
			NGT (n=3117)		411 (13)		
	Miyakoshi 2010	50 g GCT → 75 g OGTT	NGT (n=4512)	n (%)	NR (0.7)	NR	NR
			2 h IGT (8.3 mmol/L) (n=108)		NR (0)		
			1 h IGT (10.0 mmol/L) (n=66)		NR (0)		
LGA		50 g GCT	Per 1 mmol/L increment	NA	NA	OR 1.32 (1.19 to 1.46)	NR

Outcome	Study	Glucose test	Glucose threshold	Outcome unit	Outcome value	Risk (95% CI)	p-value
	Farrar 2016 HTA	75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	NA	FPG: OR 2.11 (1.73 to 2.58) 1 h: OR 1.24 (1.20 to 1.27) 2 h: OR 1.22 (1.19 to 1.25)	NR
	Donovan 2017	50 g GCT → 75 g OGTT	Normal 50 g screen (n=144,191)	n (%)	12,045 (8.2)	OR 1 (ref)	Ref
			Normal 75 g screen (n=21,248)		2270 (10.5)	OR 1.3 (1.2 to 1.4)	<0.01
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)		628 (14.2)	OR 1.7 (1.6 to 1.9)	<0.01
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767)	n (%)	Hispanic: 63 (12) Non-Hispanic white: 22 (9)	NR	NR
			Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)		Hispanic: 38 (15) NonHispanic white: 16 (14)		
	MAMMA (Berntorp 2015)	75 g OGTT	<5.7 mmol/L (n=2637)	n (%)	115 (20.1)	NR	<0.001
			5.7 to 6.4 mmol/L (n=2783)		110 (19.2)		
			6.5 to 7.2 mmol/L (n=2819)		156 (27.3)		
			Per 1 mmol/L increment (2 h glucose)	NA	NA	OR 1.09 (1.01 to 1.18)	0.028
	Jiang 2017	50 g GCT → 75 g OGTT	NGT (n=3185)	n (%)	298 (9.4)	OR **2.45 (1.46 to 4.12)**	**<0.005**
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.4 mmol/L, 2h 8.0 mmol/L (n=94)		19 (20.2)		
	Davis 2018	50 g GCT → 100 g OGTT	NGT (n=4941)	n (%)	530 (10.8)	OR 1 (ref)	Ref
			Mild hyperglycaemia (threshold unclear) (n=544)		66 (12.1)	OR 1.145 (0.87 to 1.50) AOR 0.938 (0.64 to 1.37)	0.330 0.741
			GDM/IADPSG 2010-Only (thresholds unclear) (n=181)		34 (18.9)	OR **1.932 (1.32 to 2.84)** AOR 1.466 (0.85 to 2.53)	**0.0008** 0.171
	Delibas 2018	50 g GCT → 100 g OGTT	NGT (n=316)	n (%)	9 (2.8)	NR	NR
			Single high glucose value: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=33)		1 (3.0)		
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406)	n (%)	406 (16.9)	NR	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)		115 (29.7)		
	Cheng 2009	50 g GCT →	NGT (n=13,940)	n (%)	NR (1.3)	AOR **4.28 (2.24 to 8.18)**	**<0.001**

		100 g OGTT	GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)		NR (5.1)		
	Biri 2009	50 g GCT → 100 g OGTT	NGT (n=1432)	n (%)	NR (8.0)	NR	NR
			Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326)		NR (12.0)		
			One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)		NR (14.8)		
	LIFECODES (Noor 2019)	50 g GCT	<6.7 mmol/L (120 mg/dL) (n=198)	n (%)	13 (7)	NR	NR
			6.7 to < 7.8 mmol/L (120 to <140 mg/dL) (n=47)		5 (11)		
			≥7.8 mmol/L (≥140 mg/dL) without GDM (n=32		7 (22)		
			NGT (n=4512)	n (%)	NR (6.4)	NR	Ref
Outcome	**Study**	**Glucose test**	**Glucose threshold**	**Outcome unit**	**Outcome value**	**Risk (95% CI)**	**p-value**
	Miyakoshi 2010	50 g GCT → 75 g OGTT	2 h IGT (8.3 mmol/L) (n=108)		NR (5.6)		NR **<0.05**
			1 h IGT (10.0 mmol/L) (n=66)		NR (14.6)		
Birth weight	Donovan 2017	50 g GCT → 75 g OGTT	Normal 50 g screen (n=144,191)	Mean g (SD)	3345.6 (538.5)	NR	NR
			Normal 75 g screen (n=21,248)		3345 (570.6)		
			HAPO 1.75: FPG ≥5.1 to <5.3 mmol/L,1h ≥10 to <10.6 mmol/L, 2h ≥8.5 to <9.0 mmol/L (n=4308)		3377 (605.7)		
	Beksac 2018	50 g GCT	<7.770 mmol/L (n=352)	Median g (range)	3100 (1150 to 3910)	NR	NR
			7.770 to <8.880 mmol/L (n=165)		3200 (1770 to 4150)		
			8.880 to 9.990 mmol/L (n=47)		3720 (2000 to 4280)		
			>9.990 (n=20)		3865 (2520 to 4320)		
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767)	Mean g (SD)	Hispanic: 3431 (499) 3344 (510)	NR	NR
			Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)		Hispanic: 3478 (543) Non-Hispanic white: Non-Hispanic white:		
	Delibas 2018	50 g GCT → 100 g OGTT 50 g GCT → 100 g OGTT	NGT (n=316)	Mean g (SD)	3388 (630) 3282.4 (452.8)	NR	NR
			Single high glucose value: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=33)		3290.6 (510.5)		
	Verd 2016	50 g GCT → 100 g OGTT	NGT (n=616)	Median g (range)	3272 (1995 to 4800)	NR	**0.018**
			MIGT (7.8 mmol/L to <10.6 mmol/L) (n=152)		3395 (2050 to 4390)		
	Miyakoshi 2010	50 g GCT → 75 g OGTT	NGT (n=4512)	Mean g (SD)	2957 (461)	NR	NR
			2 h IGT (8.3 mmol/L) (n=108)		2955 (439)		
			1 h IGT (10.0 mmol/L) (n=66)		3041 (401)		

Bold results are significant at p<0.05

Abbreviations: AOR, adjusted odds ratio; APR, adjusted prevalence ratio; CI, confidence interval; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; GDMFU, GDM Treatment Trial Follow-Up; HAPO, Hyperglycaemia and Adverse Pregnancy Outcomes; HTA, Health Technology Assessment; IADPSG, International Association of Diabetes and Pregnancy Study Groups; IGT, impaired glucose tolerance; MFMU, Maternal-Fetal Medicines Unit; MIGT, mild impairment of glucose tolerance; NGT, normal glucose tolerance; NR, not reported; OGTT, oral glucose tolerance test.

Respiratory distress, congenital malformation, neonatal hypoglycaemia and admission to NICU

The 1 study that reported on respiratory distress (GDMFU) found no significant difference between women with FPG thresholds of <5.1 and ≥5.1 mmol/L (4.6 % vs 5.1%, OR 1.03, 95% CI 0.34 to 3.1; p not significant).[57] Similarly, the results from the supplementary publication on the ATLANTIC-DIP cohort did not suggest any association between incremental glucose and congenital malformation (OR 0.903, 95% CI 0.309 to 2.635 for FPG to 1.095, 95% CI 0.856 to 3.960 for 1 h 75 g OGTT)[53] (Table 13).

A higher number of studies reported on neonatal hypoglycaemia (n=6) and admission to NICU (n=7) (Table 13). For neonatal hypoglycaemia, Farrar 2016 reported ORs of 1.37 (95% CI 1.20 to 1.57) and 1.13 (95% CI 1.09 to 1.18) per 1 mmol/L incremental glucose as measured by FPG and 2 h 75 g/100 g OGTT, respectively. The other 2 studies, including the GDMFU comparing glucose thresholds, did not find ORs for NGT to be lower compared to elevated glucose (0.98, 95% CI 0.27 to 3.55; p not significant[57] and 0.93, 95% CI 0.34 to 2.55[60]). The rates of neonatal hypoglycaemia were generally low, ranging from 0.4% to 4.1% for NGT and 1.2% to 6.2% for elevated glucose,[59, 61] with the exception of the MFMU Network study, where neonatal hypoglycaemia was reported in up to 21% of Hispanic women.[65]

Only 3 studies reported on measures of risk for admission to NICU and found no statistical difference between the elevated glucose and NGT groups. The GDMFU study reported an OR of 1.60 (95% CI 0.94 to 2.73; p=0.08) for elevated FPG above 5.1 mmol/L compared to <5.1 mmol/L, which changed to 1.50 (95% CI 0.78 to 2.89; p not significant) when adjusted for BMI and C-section.[57] Berggren 2011 reported an APR of 1.15 (95% CI 0.99 to 1.33) for elevated glucose compared to NGT,[69] and Cheng 2009 found no association (AOR 0.99, 95% CI 0.54 to 1.77).[60] Of the remaining 4 studies, Delibas 2018 reported that whilst admission to NICU was significantly lower in women with NGT compared to groups with reactive hyperglycaemia (9.2% vs 26.7%; p<0.05), there was no significant difference between NGT and elevated glucose.[62]

Table 13. Respiratory distress, congenital malformation, neonatal hypoglycaemia and admission to NICU

Outcome	Study	Glucose test	Glucose threshold	Outcome unit	Outcome value	Risk (95% CI)	pvalue
Respiratory distress	GDMFU (López del Val 2019)	FPG	<5.1 mmol/L (n=1193)	n (%)	54 (4.6)	OR 1.03 (0.34 to 3.21)	NS
			≥5.1 mmol/L (n=155)		8 (5.1)		
Congenital malformation	ATLANTIC-DIP (Dennedy 2012)	FPG	Per 1 mmol/L increment	NA	NA	OR 0.903 (0.309 to 2.635)	NR
		75 g OGTT 1 h glucose	Per 1 mmol/L increment			OR 1.095 (0.856 to 3.960)	
		75 g OGTT 2 h glucose	Per 1 mmol/L increment			OR 1.064 (0.770 to 1.472)	
Neonatal hypoglycaemia	Farrar 2016 HTA	50 g GCT	Per 1 mmol/L increment	NA	NA	OR 1.38 (1.00 to 1.92)	NR
		75 g and 100 g OGTT combined	Per 1 mmol/L increment	NA	NA	FPG: OR **1.37 (1.20 to 1.57)** 1 h: NR 2 h: OR **1.13 (1.09 to 1.18)**	NR
	GDMFU (López del Val 2019)	FPG	<5.1 mmol/L (n=1193)	n (%)	26 (2.2)	OR 0.98 (0.27 to 3.55),	NS
			≥5.1 mmol/L (n=155)		4 (2.6)		
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767)	n (%)	Hispanic: 84 (21) Non-Hispanic white: 25 (13)	NR	NR
			Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)		Hispanic: 30 (15) Non-Hispanic white: 13 (14)		
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	n (%)	NR (1.7)	AOR 0.93 (0.34 to 2.55)	NR
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)		NR (1.8)		
	Biri 2009	50 g GCT → 100 g OGTT	NGT (n=1432)	n (%)	NR (0.4)	NR	NR
			Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326)		NR (1.2)		
			One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)		NR (3.5)		
	Corrado 2009	50 g GCT → 100 g OGTT 50 g GCT → 100 g OGTT	One abnormal 100 g: 1 of FPG 5.3 3h 7.8 mmol/L (n=152)	n (%)	9 (6.2)	NR	0.4
			NGT (n=13,940) mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L,		26 (4.1)		
	GDMFU	FPG		n (%)	134 (11.2)	OR 1 (ref)	Ref

Admission to NICU	(López del Val 2019)		<5.1 mmol/L (n=1193) ≥5.1 mmol/L (n=155)		25 (16.5)	OR 1.60 (0.94 to 2.73) AOR[b] 1.50 (0.78 to 2.89)	0.08 NS
	MFMU Network (Berggren 2012)	50 g GCT → 100 g OGTT	Glucose intolerant: 1h 50g ≥7.5 to <11.1 mmol/L (n=767)	n (%)	Hispanic: 30 (6) Non-Hispanic white: 19 (8)	NR	NR

			Mild untreated GDM: ≥2 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=371)		Hispanic: 21 (8) Non-Hispanic white: 13 (11)		
	Delibas 2018		NGT (n=316)	n (%)	29 (9.2)	NR	NR

		50 g GCT → 100 g OGTT	Single high glucose value: 1 of FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=33)		6 (18.2)		
	Meek 2015	50 g GCT → 75 g OGTT	NGT (n=2406)	n (%)	143 (5.9)	NR	NR
			GDM/IADPSG 2010-Only: FPG 5.1 to 5.5 mmol/L, 1h ≥10.0 mmol/L, 2h <7.8 mmol/L (n=387)		22 (5.7)		
	Cheng 2009	50 g GCT → 100 g OGTT	NGT (n=13,940)	n (%)	NR (6.0)	AOR 0.99 (0.54 to 1.77)	0.91
			GDM by CC only: FPG 5.3 mmol/L, 1h 10.0 mmol/L, 2h 8.6 mmol/L, 3h 7.8 mmol/L (n=273)		NR (5.9)		
	Biri 2009	50 g GCT → 100 g OGTT	NGT (n=1432)	n (%)	NR (5.9)	NR	NR
			Abnormal 50 g (>7.8 mmol/L), normal 100 g (n=326)		NR (9.7)		
			One abnormal 100 g: 1 of 1h 10.6 mmol/L, 2h 9.2 mmol/L, 3h 8.1 mmol/L (n=142)		NR (14.8)		
	Berggren 2011	50 g GCT → 100 g OGTT	GDM by CC only (3h 7.8 mmol/L) (n=460)	n (%)	138 (30)	APR 1.15 (0.99 to 1.33)	NR
			NGT (n=3117)		804 (26)		

Abbreviations: AOR, adjusted odds ratio; APR, adjusted prevalence ratio; CC, Carpenter and Coustan; CI, confidence interval; FPG, fasting plasma glucose; GDM, gestational diabetes mellitus; GDMFU, GDM Treatment Trial Follow-Up; HTA, Health Technology Assessment; IADPSG, International Association of Diabetes and Pregnancy Study Groups; MFMU, Maternal-Fetal Medicines Unit; NGT, normal glucose tolerance; NR, not reported; OGTT, oral glucose tolerance test.

Long-term outcomes for offspring

Only 1 study, the supplementary publication for HAPO Belfast, reported on long-term outcomes for offspring. Incremental increases in FPG were significantly associated ($p<0.01$) with increased odds of BMI in the ≥85th, ≥95th and ≥99th percentile (Table 14). The effect was reduced when adjusting for BMI, age and previous GDM, becoming largely statistically non-significant (all but BMI ≥99th percentile). Neither was the effect significant for increases in glucose measured by 1 h or 2 h 75 g OGTT. A similar trend was seen for sum of skinfolds ≥90th percentile.[54]

Whilst 5 studies included in Farrar 2016 reported on longer-term outcomes in either mother or offspring, these were not included in the MA as the studies were too heterogeneous to combine. However, the results from the individual studies, as summarised in Farrar 2016, were generally consistent in that there were associations between glucose levels and sum of skinfolds along with increased body fat.

Neither of the two studies reporting on specific glucose thresholds that were comparable with NICE criteria (GDMFU and MAMMA) reported on long-term outcomes.

Table 14. BMI and sum of skinfolds in offspring aged 5 to 7 years

Outcome	Study	Glucose threshold	Outcome unit	Outcome value	Risk (95% CI)	p value
Offspring (age 5 to 7) BMI						
BMI ≥85th percentile	HAPO Belfast (Thaware 2015)	Per 1 unit rise in glucose (FPG)	NA	NA	OR **2.01 (1.37 to 2.96)** AOR 1.16 (0.76 to 1.76)	**<0.001** NR
		Per 1 unit rise in glucose (1 h glucose)			OR 1.06 (0.98 to 1.15)	NR
		Per 1 unit rise in glucose (2 h glucose)			OR 1.10 (0.99 to 1.23)	NR
BMI ≥95th percentile	HAPO Belfast (Thaware 2015)	Per 1 unit rise in glucose (FPG)	NA	NA	OR **2.37 (1.41 to 3.98)** AOR 1.34 (0.76 to 2.35)	**<0.01** NR
		Per 1 unit rise in glucose (1 h glucose)			OR 1.01 (0.91 to 1.13)	NR
		Per 1 unit rise in glucose (2 h glucose)			OR 0.99 (0.85 to 1.15)	NR
BMI ≥99th percentile	HAPO Belfast (Thaware 2015)	Per 1 unit rise in glucose (FPG)	NA	NA	OR **4.32 (2.07 to 9.04)** AOR 2.32 (1.05 to 5.13)	**<0.001** NR
		Per 1 unit rise in glucose (1 h glucose)			OR 1.06 (0.90 to 1.24)	NR
		Per 1 unit rise in glucose (2 h glucose)			OR 0.94 (0.75 to 1.18)	NR
Offspring (age 5 to 7) sum of skinfolds ≥90$_{th}$ percentile	HAPO Belfast (Thaware 2015)	Per 1 unit rise in glucose (FPG)	NA	NA	OR **2.48 (1.44 to 4.26)** AOR 1.61 (0.90 to 2.89)	**<0.01** NR
		Per 1 unit rise in glucose (1 h glucose)			OR 1.02 (0.91 to 1.14)	NR
		Per 1 unit rise in glucose (2 h glucose)			OR 0.99 (0.84 to 1.16)	NR

Bold results are significant at $p<0.05$

Abbreviations: AOR, adjusted odds ratio; CI, confidence interval; FPG, fasting plasma glucose; HAPO, Hyperglycaemia and Adverse Pregnancy Outcomes; normal glucose tolerance; NR, not reported; OR, odds ratio

Conclusions

The aim of this question was to identify associations between incremental increases in glucose levels and the risks of adverse pregnancy or neonatal outcomes in low-risk women, in order to distinguish a threshold at which risk is increased above normal but still considered to be low risk based on the current NICE guidance (presence of risk factors; 75 g OGTT: FPG, ≥5.6 mmol/L; 2h OGTT, ≥7.8 mmol/L).

High-to-moderate quality evidence, including a high quality SLR that also included hightomoderate quality evidence, was identified for associations between elevated glucose and pregnancy, neonatal and long-term outcomes. A summary of the direction of associations between NGT and elevated glucose groups is presented in Table 15. Overall, clear and consistent associations between risk and increased glucose were not identified for the majority of outcomes. The exceptions were C-section, induction of labour, macrosomia and LGA (and pre-eclampsia and hypertension in some studies), where increased glucose (compared with NGT) was associated with increased risk.

However, very limited evidence was identified that would allow for the characterisation of a specific glucose threshold compared with that of the NICE guidance. Only 2 studies, MAMMA and GDMFU, investigated single glucose thresholds that were elevated from normal but lower than 7.8 mmol/L (the current NICE cut-off) which was comparable to that of NICE (e.g. measured using 75 g OGTT or FPG). The remaining studies either a) used criteria for "elevated" that could have included a number of abnormal values from different measures (e.g. abnormal on either FPG, 1 h, 2 h or 3 h 100 g OGTT to qualify as "elevated glucose") and testing was heterogenous across different studies (therefore not possible to distinguish a single threshold); and/or b) used a different test to that recommended by NICE (e.g. used 100 g whereas NICE recommends 75 g OGTT) so the elevated values would not be comparable. Simply due to the lack of evidence, it was not possible to identify a threshold at which risk for particular outcomes is substantially elevated. However, it was noted that the MAMMA and GDMFU studies often reported significantly higher rates of adverse outcomes in their elevated glucose groups, and so these thresholds (GDMFU: 5.1 mmol/L; MAMMA 5.7 mmol/L) could be considered as a starting point for an association with increased risk of adverse outcomes. This is particularly apparent in macrosomia and LGA, where the association between elevated glucose and the outcome was also supported by other included studies, even though their populations were not clearly below the NICE threshold. (Table 15).

Table 15. Summary of the number of studies and direction of evidence for each outcome

Outcome	Number of studies reporting	Direction of evidence	Possible to identify threshold for increased risk?
Pregnancy outcomes	**16**		

Gestational age	10	No clear association	No – only reported by 1 study, **MAMMA** (but significantly shorter gestational age)
Pre-term birth	10	No clear association	No – only reported by 1 study, **MAMMA** (but significantly higher pre-term birth)
Pre-eclampsia	8	Association between ↑ glucose and ↑ risk in 2 studies (including Farrar 2016); no clear association in 3 studies; NR in 3 studies	No – only reported by 1 study, **GDMFU** (no significant difference)
Hypertension	6	Association between ↑ glucose and ↑ risk in 3 studies; no clear association in 2 studies; NR in 1 study	No – not reported
Neonatal outcomes	**19**		
Perinatal mortality/stillbirth	4	No clear association	No – not reported
C-section	12	Association between ↑ glucose and ↑ risk	No – not reported
Induction of labour	3	Association between ↑ glucose and ↑ risk	No – not reported
Shoulder dystocia	5	Association between ↑ glucose and ↑ risk in 2 studies (including Farrar 2016); no clear association in 3 studies	No – not reported
Lacerations	3	No clear association	No – not reported
Trauma during vaginal delivery	2	Association between ↑ glucose and ↑ risk in 1 study; no clear association in 1 study	No – only reported by 1 study, **GDMFU** (but significantly higher birth trauma)
Macrosomia	11	Association between ↑ glucose and ↑ risk	No – only reported by 1 study, **GDMFU** (but significantly higher macrosomia)
LGA	12	Association between ↑ glucose and ↑ risk	No – only reported by 1 study, **MAMMA** (but significantly higher LGA)
NICU admission	7	No clear association	No – only reported by 1 study, **GDMFU** (but significantly higher NICU admission)
Respiratory distress	1	No clear association	No – only reported by 1 study, **GDMFU** (no significant difference)
Congenital malformation	1	No clear association	No – not reported
Neonatal hypoglycaemia	6	Association between ↑ glucose and ↑ risk in 1 study (Farrar 2016); no clear association in 2 studies	No – only reported by 1 study, **GDMFU** (no significant difference)
Long-term outcomes	**1**		

BMI	1	Association between ↑ glucose and ↑ risk for FPG; no clear association for 1 h or 2 h 75 g OGTT	No – not reported
Sum of skinfolds	1	Association between ↑ glucose and ↑ risk for FPG; no clear association for 1 h or 2 h 75 g OGTT	No – not reported

Furthermore, whilst Farrar 2016 was a useful source that performed a MA and reported the increase in risk per 1 mmol/L unit increase in glucose in a dose-response manner, it was not possible to use this to identify the threshold at which risk is substantially greater. Indeed, Farrar 2016 reported that their analysis found that the odds of adverse outcomes increased linearly with glucose levels, suggestive of a continuum risk across glucose levels, and no clear threshold that can define elevated glucose. This suggests that even performing a MA, as in Farrar 2016, would not provide insight into a specific glucose cut-off point.

It may be that a standardised "threshold" for elevated glucose would need to be based on meeting specific diagnostic criteria or on number of abnormal glucose values rather than a single numerical threshold. Within this work, it was not possible to identify a specific set of criteria that conferred consistently higher risk for any particular outcome within populations that would be considered to be low-risk according to the current NICE criteria, due to a wide variety of criteria being used across studies and inconsistent results in associations.

Summary of Findings Relevant to Criterion 1: Criterion not met*

Quantity: A large volume of evidence was identified overall, including 1 SLR with an MA of 38 publications, and 17 studies found through database searches. Evidence was identified for a large number of specific pregnancy and neonatal outcomes, with many reported in at least 5 studies. Studies reported a consistent association between elevated glucose and increased risk of C-section (12 studies), induction of labour (3 studies), macrosomia (11 studies) and LGA (12 studies). Data for all other outcomes was either limited or there was no clear association between elevated glucose and risk level. Furthermore, a very limited quantity of evidence was identified to address the question of identifying a specific threshold for elevated glucose at which there is an increase in risk substantial enough that would justify population screening within low-risk women that would differ from the current NICE guidance. This was only possible to explore in 2 studies that reported single thresholds using tests that could be compared to the current NICE recommendations.[57, 58]

Quality: The quality of the Farrar 2016 SLR was judged to be high, especially with results where inter-study heterogeneity was considered. Studies identified from the database searches were all judged to be at moderate (n=9, including MAMMA and GDMFU) or serious (n=8) risk of bias. A main concern was in confounding through maternal risk factors that were not adjusted for and may have therefore influenced the risk of outcomes concomitantly to elevated glucose. There was also a concern that outcome assessors were not blinded to women's glycaemic status, which in some cases may have introduced systematic errors in the measurement of the less objective outcomes. For such studies, concerns for bias were lower. Another less likely (but nonetheless noted) source of bias was lack of a published protocol or SAP that would allow for an easier assessment of the risk of bias in reporting of outcomes. The concerns identified for the database studies were the same as the main concerns described by Farrar 2016 for the studies included in their SLR. For all other domains, the majority of studies were judged to be at low risk of bias.

Applicability: The main concern for applicability to a UK setting arises from the inclusion of 7 non-EEA/OECD countries in the Farrar 2016 SLR. However, this was a low

* **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

proportion of all included studies. Otherwise, all studies included from the database searches were conducted in EEA or OECD countries and are considered to be applicable to the UK setting. The lack of applicability to the current NICE guidance in terms of test used and number of abnormal values limited the majority of studies from feasibility to identify a threshold of elevated risk.

Consistency: The only outcomes where results were consistent were for C-section, induction of labour, macrosomia and LGA, where elevated glucose was associated with an increased risk of the outcome in all studies. Results for all other outcomes were inconsistent. Key inconsistencies in methodologies of included studies were different glucose tests and different criteria or thresholds for elevated glucose. This makes it difficult to determine a common threshold constituting "elevated glucose" from normal levels. Farrar 2016 provided a satisfactory discussion of the heterogeneity of its included studies, noting that whilst heterogeneity was considerable for studies reporting on macrosomia and LGA, the trends were reasonably consistent across different studies, and that there was no evidence that trend in risk with glucose level is different for different glucose tests.

Conclusions: Moderate-to-high quality evidence for a wide number of pregnancy and neonatal outcomes was identified in this rapid review. The evidence was judged to be broadly applicable to the UK clinical setting. However, while the review identified clear associations from a large volume of evidence between elevated glucose and increased risk of C-section, induction of labour, macrosomia and LGA, results for other outcomes were inconsistent. Macrosomia and LGA were also significantly increased in women who would not currently be identified as at risk by the NICE guideline, but neither C-section nor induction of labour was reported by either study investigating low risk women. Furthermore, a clear glucose threshold for increased risk could not be identified for any outcome, mostly due to the limited evidence on single thresholds. This is reflective of the findings from Farrar 2016 and the HAPO study that there is a continuum of risk across increasing glucose levels and no clear cut-off point. On this basis, Criterion 1 was judged to be not met.

Criterion 4 — There should be a simple, safe, precise and validated screening test.

Question 2 – What are the most effective screening tests or strategies to identify low risk women at risk of hyperglycaemia in pregnancy or GDM?

The rapid review conducted for the UK NSC by Waugh and colleagues in 2010 identified studies on screening for GDM, the majority of which compared screening tests such as FPG and the 50 g GCT against OGTT. A key conclusion of this review was that most of the

identified studies used screening tests to identify the presence, or not, of GDM based on various forms of the OGTT but were not used to identidy elevated risk in an unselected population of pregnant women.[5]

Eligibility for inclusion in the review

This review searched for diagnostic test accuracy, cross-sectional, cohort and case-control studies, as well as SLRs and MAs of those, published since January 2009. Studies were included if the population comprised unselected pregnant women without pre-existing diabetes and specific risk factors (i.e. women who would receive a test for GDM in the current NICE pathway). Screening tests of interest included, but were not limited to, tests measuring maternal glucose, maternal history or risk factors, and/or predictive biomarkers to detect GDM. Studies of tests aiming to predict the risk of developing GDM at a further point in the pregnancy were not eligible. Studies were not excluded based on the reference standard used in the study. Outcome measures of interest for question 2 were measures of screening accuracy (e.g. area under the curve [AUC], sensitivity, specificity, positive predictive value [PPV] and negative predictive value [NPV]). Studies were restricted geographically to OECD or EEA countries, excluding Mexico and South Korea. Full details of eligibility criteria are presented in Table 4.

A high quality SLR conducted as part of an HTA by Farrar and colleagues (2016), was identified and formed the evidence base for question 2 (Q2). The aim of Farrar 2016 was to evaluate the performance of risk factors in identifying women with GDM, and the eligibility criteria were closely aligned to the eligibility criteria of this rapid review for Q2. However, there were some differences. For example, the search strategy used in Farrar 2016 did not include terms for biomarker tests, which was accounted for by adding search terms for these outcomes. Furthermore, Farrar 2016 did not use a geographic limit and included 6 studies from non-OECD/EEA countries (China, India, Iran, Malaysia and Thailand). Separate results excluding these countries from the authors summary were not available, therefore this has been noted as a limitation when considering the generalisability of the results to a UK setting.

In the UK, women are diagnosed with GDM either with an FPG ≥5.6 mmol/L or a 2 h 75 g OGTT test ≥7.8mmol/L. The latter (75 g OGTT) is recommended as the test of choice for women with risk factors for GDM by the NICE NG3 guidelines. While the test is considered accurate in diagnosing the condition, it includes a glucose loading step, which may be harmful in itself to those with poor glucose tolerance. As such, the test itself is a possible risk to those it aims to diagnose and women with risk factors are potentially at a greater risk of being harmed by the test. It is therefore relevant to identify whether any other tests, especially those not involving glucose loading could be comparable to the OGTT in the accuracy of GDM detection in the population of low-risk pregnant women.

Description of the evidence

A total of 25 publications on 21 studies were initially included from the database searches, with no additional publications found through hand-searches. Due to the high number studies identified as relevant, 7 case-control studies were ultimately not selected for extraction, as this study design is generally of lower methodological quality and at a higher risk of bias and confounding. Figure 1 (Appendix 1) depicts the flow of the included records using a PRISMA flow diagram.

Ultimately, in addition to 2 publications on the Farrar 2016 SLR, 16 articles on 13 unique studies were selected for extraction for Q2. The smallest study recruited 202 pregnant women,[74] and the largest study recruited 16,537 women.[4]

None of the relevant studies were conducted in the UK; 2 studies were conducted in the US and 2 in Japan, and 1 in each of: Australia, Belgium, Canada, The Netherlands, Poland, Spain, Sweden, Switzerland and Turkey. The Farrar 2016 SLR included studies from 18 different countries plus an analysis of individual patient data from the Born in Bradford cohort (from the UK) and the ATLANTIC-DIP cohort (from the Republic of Ireland). The populations often included women with risk factors, thought only some of these were the same risk factors as those listed by the NICE guidance as high risk for GDM.

Seven studies were of a prospective cohort design (including 1 model development study), 4 were retrospective analyses,[75-78] and 2 cross-sectional studies.

Nearly half of the identified studies (5 and the Farrar 2016 SLR and IPD analyses) evaluated 1 or more combinations of maternal risk factors for identifying women with GDM, with Temming 2016 and Saeedi 2018 exploring the accuracy of using risk factors (including history of previous GDM, obesity, history of prior macrosomic/LGA infant, first-degree relative with diabetes mellitus) alongside glucose levels.[78, 79] Five studies evaluated the use of glucose tests for identification of GDM, including the 1 h GCT[74, 77, 80] and the OGTT,[75, 81] although the thresholds for classification of a positive screen result varied. Finally, 3 studies evaluated the use of biomarkers in identifying women with GDM, including HbA1c,[82] fructosamine,[83] and various lipid and apolipoprotein markers.[84]

The reference standard used by the majority of studies was a variation of or a continuation onto (for those investigating the GCT as part of the index test) the 2-step screen for GDM.

In others, it was based either on a single glucose challenge or tolerance test. The included studies are summarised in Table 16.

Table 16. Summary of studies included for question 2

Index test	Comparator	Study	Country	Index test threshold(s)	Reference standard(s) and threshold	Outcomes
BEDIP-N[80, 85-87]	Belgium	GCT	7.2 mmol/L	Two-step 50 g GCT and 75 g OGTT using the 2013 WHO criteria	-	Abnormal GCTs, sensitivity, specificity, LR+, LR-, positive post-test probability, negative posttest probability
Pawelec 2009[74]	Poland	1-h 50 g GCT measured by finger capillary glucometer	>140 mg/dL	2 h 75 g OGTT (cut-off 155 mg/dL or 8.6 mmol/L)	Venous blood glucose >140 mg/dL	Specificity, PPV

Ryser Ruetschi 2016[81]	Switzerland	FPB	≤4.4 mmol/L 4.4 to 5.1 mmol/L ≥5.1 mmol/L	1 h and 2 h OGTT using the IADPSG criteria	-	Sensitivity, specificity, women correctly diagnosed with GDM, women avoiding glucose overload
Gringas 2018 (Project Viva)[83]	United States	Fructosamine	50th (222 µmol/L), 75th (256 µmol/L) and 95th (312 µmol/L) percentiles	Two-step 1 h 50 g GCT followed by 3-hour 100 g OGTT in screen positive women	-	Sensitivity, specificity, PPV, NPV
Iimura 2015[84]	Japan	Lipid and apolipoprotein markers	NR	Two-step 1 h 50 g GCT followed by 75 g OGTT in screen positive women	-	AUC
Khalafallah 2016[82]	Australia	HbA1c	Varied between 4.6 and 10%	One-step 75 g 2 h OGTT in line with the ADIPS consensus guidelines	-	Sensitivity, specificity, PPV, NPV
Kosus 2012[75]	Turkey	3-hour 100 g OGTT	Various for each timepoint	Two-step 1 h 50 g GCT followed by 3-hour 100 g OGTT in screen positive women	-	Sensitivity, specificity, AUC
Maesa 2018[76]	Spain	FPB	Varied between 55 and 80 mg/dL	Two-step 1 h 50 g GCT followed by 100 g OGTT in screen positive women	-	Sensitivity, specificity, PPV, NPV, LR+, LR-
Ohara 2016[77]	Japan	50 g GCT	7.8 mmol/L	75 g OGTT, using universal criteria established by the IADPSG criteria (5.1 mmol/L), the 1-h cut-off value (10.0 mmol/L), or the 2-h cut-off value (8.5 mmol/L)	-	PPV
			mmol/L	accordance with 1980 WHO criteria		NPV, AUC
Temming 2016[78]	United States	1-hour 50 g GCT with and without risk factors	>140 mg/dL (elevated) >180 mg/Dl (extremely elevated)	3 h 100 g OGTT. GDM was diagnosed by having 2 or more abnormal values using NDDG criteria (fasting ≥105 mg/dL, 1-hour ≥190 mg/dL, 2-hour ≥165 mg/dL, 3hour ≥145 mg/dL) or using more stringent CC criteria (fasting ≥95 mg/dL, 1-hour ≥180 mg/dL, 2-hour ≥155 mg/dL, 3-hour ≥140 mg/dL	-	Sensitivity, specificity, AUC, PPV

Index test	Comparator	Study	Country	Index test threshold(s)	Reference standard(s) and threshold	Outcomes
Saeedi 2018[79]	Sweden	Four clinical FPB	Varied at 4.0 to 4.5	One-step 75 g 2 h OGTT in	-	Sensitivity, specificity, PPV,

Theriault 2014[88]	Canada	risk factor models assessed at 24–28 weeks	-	50 g GCT in all women followed by 75 g OGTT if GCT between 7.8–10.2 mmol/L	-		Sensitivity, specificity, PPV, NPV, AUC
Van Leeuwen 2010[89]	Netherlands	Universal testing with OGTT	-	Two-step plasma glucose and 50 g GCT followed by 75 g 2 h OGTT in those with abnormal value on the first tests, according to the WHO criteria	Diagnostic testing if the probability of GDM is ≥2.0% or ≥4.0%		Sensitivity, specificity, OGTT to diagnose one case of GDM, n
Farrar 2016 (SLR)[90]	Various; IPD from mmol/L)	Risk factor models	-	Modified WHO 1988 criteria (FPG >6.1 mmol/L, 2-hour post-load	-	UK and ROI	glucose >7.8

Abbreviations: AUC: area under the curve, CC: Carpenter and Coustan, FPG: fasting plasma glucose, GCT: glucose challenge test, GDM: gestational diabetes; IADPSG, International Association of Diabetes and Pregnancy Study Groups; LR: likelihood ratio, NDDG, National Diabetes Data Group; NPV: negative predictive value, OGTT: oral glucose tolerance test, PPV: positive predictive values, ROI: Republic of Ireland, WHO: World Health Organisation

Quality assessment (Q2)

Farrar 2016

The quality of the included Farrar 2016 SLR was appraised using the AMSTAR 2 checklist; a summary is given here and in Table 101 (Appendix 4).

The quality of Farrar 2016 was high overall, including clear objectives and eligibility criteria, a comprehensive search strategy and robust methodology (dual review). The results were clearly reported including a detailed discussion of the characteristics of included studies (no MA was conducted due to heterogeneity across the included studies). The quality of the included studies was not assessed, justified by the authors due to lack of an appropriate tool. Therefore, there is uncertainty around the quality of included studies, and it is noted that an attempt to modify the QUADAS-2 checklist could have been made. Farrar 2016 also failed to report on the source of funding in the included studies; nevertheless, this was also not expected to increase the risk of bias as most studies appear to have been conducted in academic environments.

It should be noted that the eligibility criteria for Farrar 2016 differed from this rapid review in that studies from any country were eligible, rather than being limited to OECD or EEA countries. Six studies from non-OECD/EEA countries were included in the authors' summary, which is noted as a limitation to the applicability of the Farrar 2016 review.

Other included studies from database searches

The quality of the 11 included studies that assessed screening tests for GDM was appraised using an adapted QUADAS-2 checklist (Appendix 4). The quality of the 2 studies that assessed models was appraised using an adapted PROBAST tool checklist (Appendix 4). A summary of the risk of bias and applicability to the UK setting is presented in Table 17 and Table 18, and the full appraisals are presented in Appendix 4. Overall, risk of bias was judged to be low in the majority of studies for participant selection and reference standard but high/unclear in the majority for index tests and participant flow.

Table 17. Summary of QUADAS-2 assessments for GDM screening studies

Question	Khalafallah 2016[82]	BEDIP-N[80]	Gringas 2018[83]	Iimura 2015[84]	Kosus 2012[75]	Maesa 2018[76]	Ohara 2106[77]	Pawelec 2009[74]	Ryser Ruetschi 2016	Saeedi 2018[79]	Temming 2016[78]
PARTICIPANT SELECTION											

Risk of bias	Low	Low	Low	Low	High	Low	Unclear	Unclear	Low	Low	Low
Concern about applicability	Low	Low	Unclear	Low	Unclear	Low	Unclear	Unclear	Low	Low	Unclear
INDEX TESTS											
Risk of bias	High	Unclear	Unclear	High	High	High	Low	Low	Low	High	Low
Concern about applicability	Low	High	High	Low	Unclear	Low	Low	Low	Low	Low	Low
REFERENCE STANDARD											
Risk of bias	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low
Concern about applicability	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
PARTICIPANT FLOW											
Risk of bias	Low	Low	High	Unclear	High	High	High	High	Low	Low	High

Participant selection

The risk of selection bias was judged as low in 8 of 11 studies due to consecutive enrolment, and for avoiding a case-control study design. Kosus 2018, a retrospective cohort study, was considered to be at a high risk of selection bias due to exclusion of women with family history or previous complications such as history of GDM, congenital anomalies, unexplained fetal loss, hypertension or stillbirth.[75] Excluding women with risk factors that would not be covered by the NICE pathway may limit the applicability (therefore judged as unclear risk of bias) of the sample to the general UK population. Two studies were judged to be at an unclear risk of bias and unclear for their applicability to the UK population, due to poor reporting of the recruitment methods and eligibility criteria.[74, 77] Another 2 studies also had unclear applicability because the makeup of their population may be somewhat different to that in the UK: one had a high proportion of African American women,[78] and the other included singleton pregnancies in Japan.[84]

Index tests

Five studies were at high risk of bias for how the index test was conducted, as the thresholds for classifying a test result as positive were not prespecified and it was either unclear when the measures were taken or whether they were interpreted without the knowledge of the reference standard.[75, 76, 79, 80, 82] Two studies were at an unclear risk of bias; in the Gingras 2018 study, it was not reported whether study assessors were blind to glucose results when testing fructosamine levels from samples previously collected in 1999–2002.[83] The risk of bias was also unclear in Iimura 2015 because insufficient information was available.[84]

There was a high concern about applicability in 2 studies because the tests examined were not currently used in the UK clinical practice (fructosamine, lipid biomarkers).[83, 84] Index tests used in all other studies were judged applicable as already used in the UK clinical practice either as part of diabetes or GDM targeted testing (in at-risk individuals).

Reference standard

Ten out of 11 included studies were at low risk of bias for conduct and interpretation of the reference standard; in the majority of studies, published diagnostic criteria were used including the National Diabetes Data Group (NDDG), CC, IADPSG criteria and there were no concerns about the correct diagnosis of women with GDM. In Saeedi 2018, the authors noted some uncertainty around the validity of using capillary instead of venous blood and therefore the risk of bias in the application of the reference standard was judged as unclear in that study.[79] While healthcare providers and participants were blinded to the results of the index test in only 1 study,[80] the use of published, objective criteria reduced the risk of bias in the interpretation of the reference standard in the other studies.

There were low concerns about applicability to the UK setting in all studies, due to the use of established diagnostic criteria, which bear similarity to current UK practice guidelines.

Participant flow

Six studies were at high risk of bias for this domain. This was mostly because only screen-positive women were tested with the reference standard. While this approach is the accepted 'norm' in clinical practice and offering the reference standard (OGTT being the only currently available reference standard) to screen-negative women might be considered unethical due to the risks associated with glucose loading, it increases the risk of partial verification bias and can lead to overestimation of sensitivity or underestimation of specificity (if screen-negative women are not confirmed as true negatives). In addition, the Temming 2009 and Ohara 2016 studies excluded women who were lost to follow-up or have not completed the 3 h OGTT, or those with hyperemesis gravidarum respectively. Meanwhile in Maesa 2018 and Iimura 2015, the time interval between the index test and the reference standard was unclear, increasing the risk that interventions or changes to lifestyle may have occurred during this timeframe, which could have affected the result of the reference standard.[76-78, 84] Furthermore, up to 50% of women in Iimura 2015 had missing biochemical or lipoprotein data.[84]

Khalafallah 2016 was at an unclear risk of bias, due to uncertainty around all women receiving the reference standard or being included in the analyses.[82] In the remaining studies, all or almost all participants were included in the analyses, and all screened participants received the same reference standard.

Predictive model studies

The review identified 2 predictive studies, van Leeuwen 2010 and Theriault 2014, the quality of which was assessed with the PROBAST checklist and is summarised in Table 18.

Table 18. Summary of PROBAST assessments for GDM screening studies

Question	Van Leeuwen 2010[89]	Theriault 2014[88]
Type of prediction study	Development only	Validation only
PARTICIPANT SELECTION		
Risk of bias	Low	Low
Concern about applicability	Low	Low
PREDICTORS		
Risk of bias	Low	Low
Concern about applicability	Low	Low
OUTCOME		
Risk of bias	Low	High
Concern about applicability	Low	High
ANALYSIS		
Risk of bias	High	High
Overall assessment		
Risk of bias	High	High
Concern for applicability	Low	High

Participant selection

Both studies used appropriate inclusion/exclusion criteria and data sources and were thus judged to be at a low risk of selection bias. There were also no concerns about the applicability of the included populations to the UK.

Predictors

There were no concerns about bias or applicability due to predictors in either of the 2 studies. All predictors appeared to have been assessed in the same way for all women, before the knowledge of the outcome and would be available by the point in pregnancy when the model needs be used.

Outcomes

The risk of bias and applicability due to outcomes were judged as low and of no concerns in Van Leeuwen 2010. However, risk of bias and applicability were high and of concern in Theriault 2014 as GDM was not diagnosed in the same way in all women; some received glucose tests but others were only deemed to have GDM as they had used insulin in pregnancy.

Analysis

Both studies were judged to be at a high risk of bias in the results because of the analyses' methodology. In Van Leeuwen 2010, this is because the sample was small and it was unclear how many women were eventually enrolled; it was likely that performance of the model was not measured appropriately or that overfitting was not accounted for.[89] The Theriault 2014 model was judged at the high risk of bias because of how the authors handled missing data, where some women were included in the analysis despite missing test results.[88]

Results

Key results for each of the screening tests are presented in Table 19, Table 20, Table 21, Table 22 and Table 23. Full details of the included studies and their results can be found in Appendix 3.

Oral glucose tolerance test

Two studies evaluated the OGTT as a screening test for GDM (Table 19). Kosus 2012 (n=808) investigated the fasting, 1-hour, 2-hour and 3-hour 100 g OGTTs at various cut-off levels, with the objective of identifying the optimal cut-off for high sensitivity and specificity using the CC criteria as a reference standard.[75] For all OGTTs, the use of lower, conservative cut-off levels resulted in high sensitivities, reaching 100% using a cut-off of 145.5 mg/dL with a 1 h OGTT, but this was at the expense of specificity (37.3%). Based on the selected cut-offs, the 2 h OGTT was found to have highest screening accuracy (sensitivity 88.1, specificity 87.6, AUC 0.911), compared with FPG (sensitivity 82.1%, specificity 52.2%, AUC 0.752), 1 h (sensitivity 83.6%, specificity 80.1%, AUC 0.894) and 3 h (sensitivity 74.6, specificity 60.2, AUC 0.782) OGTTs.[75]

Ryser Ruetschi 2016 (n=2298) aimed to evaluate how the fasting measurement of blood glucose alone, prior to loading for the OGTT, could reduce the number of women requiring further testing with a glucose load, at various FPG cut-off values.[81] Unsurprisingly, sensitivity was highest (96%) and specificity lowest (25.3%) at the most conservative cut-off of 4.0 mmol/L, with a sensitivity of 47.4% and specificity of 100% at a 5.1 mmol/L cut-off. Ryser Ruetschi 2016 further evaluated 2 screening strategies: 1) a strategy of stopping the test, avoiding glucose loading and further glycemia, if fasting glucose was <4.4 or ≥5.1 mmol/L; and 2) excluding women with a fasting glycaemia greater than 5.1 mmol/L. The first strategy was successful in avoiding loading in 69% of women and achieved a sensitivity of 78.5% (95% CI 73.1 to 83.2). The second strategy resulted in a sensitivity of only 59.1% in the remaining population. Specificities for the 2 strategies were not reported.[81]

Table 19. Screening for GDM using an oral glucose tolerance test

			Test accuracy						
Study	**Test**	**Threshold/cut-off**	**Sens. (95% CI)**	**Spec. (95% CI)**	**PPV (95% CI)**	**NPV (95% CI)**	**LR+**	**LR-**	**AUC (95% CI)**
	Fasting 100 g OGTT	82.5 mg/dL	82.1	52.2	NR	NR	NR	NR	0.752 (0.678 to 0.825)
Kosus 2012[75]	1-hour 100 g OGTT	171.5 mg/dL	83.6	80.1	NR	NR	NR	NR	0.894 (0.854 to 0.934)
	2-hour 100 g OGTT	151.5 mg/dL	88.1	87.6	NR	NR	NR	NR	0.911 (0.868 to 0.954)
	3-hour 100 g OGTT	111.5 mg/dL	74.6	60.2	NR	NR	NR	NR	0.782 (0.708 to 0.857)
Ryser Ruetschi 2016[81]	OGTT	4.0 mmol/L	96.0	25.3	NR	NR	NR	NR	NR
		4.2 mmol/L	88.8	47.9	NR	NR	NR	NR	NR
		4.4 mmol/L	78.5	69	NR	NR	NR	NR	NR
		4.6 mmol/L	67.7	84.3	NR	NR	NR	NR	NR
		5.1 mmol/L	47.4	100	NR	NR	NR	NR	NR

Abbreviations: AUC: area under the curve, CI, confidence interval; LR: likelihood ratio, NPV: negative predictive value, OGTT: oral glucose tolerance test, PPV: positive predictive values

Glucose challenge test

The accuracy of the GCT as a screening test on its own was evaluated in 4 studies (Table 20). BEDIP-N and Temming 2016 both evaluated the GCT test at various cut-offs; in BEDIP-N the GCT was done as part of a universal 2-step screening strategy using the 2013 WHO criteria.[80] Temming 2016 checked how many women would be diagnosed with GDM with just the GCT vs a 2 h 100 g OGTT as per the NDDG or the CC criteria.[78]

Results of BEDIP-N (n=1884) showed that ≥130 mg/dL (7.2 mmol/L) was the optimal cutoff in terms of balancing sensitivity (72.4%, 95% CI 66.1 to 78.1) and specificity (70.2%, 95% CI 67.9 to 72.4), but was described as having moderate accuracy. A threshold of <7.2 mmol/l was not recommended, and while the test can achieve a higher sensitivity of up to 82% (6.7 mmol/L cut-off), this would be at the expense of specificity (56%).[80] Temming 2016 (n=753) only used thresholds of 160 mg/dL and 180 mg/dL and found sensitivity at either to be worse than in BEDIP-N, regardless of the criteria used in the reference standard.[78] Specificity was better than in BEDIP-N with the higher diagnostic threshold (180 mg/dL, equivalent to 10 mmol/L) with both NDDG (92.2%) and CC (93.2%) criteria, but this was at a significantly reduced sensitivity. Interestingly, at the lower threshold of 160 mg/dL (equivalent to 8.9 mmol/L, still higher than the 7.8 mmol/L threshold used by BEDIP-N) the specificity appeared higher in BEDIP-N than Temming 2016, though this may be within the margin of variability or due to the use of different diagnostic criteria for the reference standard. Conversely, a very high sensitivity was shown by capillary blood sampling at the 140 mg/dL cut-off in Pawelec 2009 (n=202) (98.5%), though the specificity of approach was

at only 66.7%.[74] The study by Ohara 2016 (n=2079) only reported the PPV for the 50 g GCT, which was 42.8%.[77]

Based on the limited evidence it appears that the higher GCT thresholds can decrease the number of women who would need an OGTT but then test negative for GDM, but this results in many women who do develop GDM to be missed and not receive treatment. If a lower threshold for GCT is used, the situation is reversed in that fewer GDM cases can be detected but more women undergo OGTTs.

Table 20. Screening for GDM using a glucose challenge test

Study	Screening test	Threshold for a positive result	Reference standard	Sens. (95% CI)	Spec. (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR– (95% CI)	AUC
BEDIP-N[80, 85]	GCT	7.8 mmol/L (≥140 mg/dL) 7.5 mmol/L(≥135 mg/dL) 7.2 mmol/L (≥130 mg/dL) 6.9 mmol/L (≥125 mg/dL) 6.7 mmol/L (≥120 mg/dL)	75 g OGTT with 2013 WHO criteria (FPG ≥5.1 mmol/L, 1 h glycaemia, ≥10.0 mmol/L, 2 h glycaemia, ≥8.5 mmol/L, diagnosis of GDM if ≥1 value is abnormal)	59.6 (53.0–66.1)	81.0 (79.0–82.9)	NR	NR	3.1 (2.7–3.6)	0.50 (0.42–0.58)	NR
				66.2 (59.7–72.3)	76.1 (73.9–78.1)	NR	NR	2.8 (2.4–3.1)	0.44 (0.37–0.53)	NR
				72.4 (66.1–78.1)	70.2 (67.9–72.4)	NR	NR	2.4 (2.2–2.7)	0.39 (0.32–0.49)	NR
				77.6 (71.7–82.9)	64.2 (61.8–66.5)	NR	NR	2.2 (2.0–2.4)	0.35 (0.27–0.45)	NR
				82.0 (76.4–86.8)	56.0 (53.5–58.4)	NR	NR	1.9 (1.7–2.0)	0.32 (0.24–0.43)	NR
Temming 2016[78]	GCT	≥160 mg/dL (8.9 mmol/L)	3 h 100 g OGTT, NDDG criteria	65.5 (57.7–72.7)	70.2 (66.4–73.9)	38.2 (32.5 to 44.1)	NR	NR	NR	0.678 (0.638 to 0.719)
		≥180 mg/dL (10 mmol/L)		30.3 (23.4 to 37.9)	92.2 (89.7 to 94.2)	52.1 (41.6 to 62.4)	NR	NR	NR	0.612 (0.576 to 0.649)
		≥160 mg/dL (8.9 mmol/L)	3 h 100 g OGTT, CC criteria	58.4 (52.0–64.6)	72.8 (68.6–76.6)	51.6 (45.6 to 57.5)	NR	NR	NR	0.656 (0.62 to 0.692)
		≥180 mg/dL (10 mmol/L)		24.8 (19.6 to 30.6)	93.2 (90.7 to 95.3)	64.6 (54.2 to 74.1)	NR	NR	NR	0.590 (0.561 to 0.619)
Ohara 2016[77]	GCT	7.8 mmol/L	75 g OGTT, IADPSG criteria for GDM	NR	NR	42.8	NR	NR	NR	NR
Pawelec 2009[74]	GCT	Finger capillary blood sample using glucometer (>140 mg/dL cut-off)	2 h 75 g OGTT (cut-off 155 mg/dL or 8.6 mmol/L)	98.5	66.7	NR	NR	NR	NR	NR

Abbreviations: AUC: area under the curve; CC, Carpenter and Coustan; CI, confidence interval; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; LR: likelihood ratio, NDDG, National Diabetes Data Group; NPV: negative predictive value; NR, not reported; OGTT: oral glucose tolerance test, PPV: positive predictive values; WHO, World Health Organization

Fasting plasma glucose

FPG was investigated as a test to diagnose GDM in 1 study: Maesa 2018 (n=6573) (Table 21).[76] Maesa 2018 investigated the number of women in whom GDM could be ruled out depending on various thresholds of FPG (though not explicitly reported in the publication, it was inferred that FPG was measured). As expected, the lower the threshold, the higher the sensitivity of the test (up to 97.8%), but the lower the specificity – at 55 mg/dL FPG could only rule out 1.28% of women, which meant that almost all women had to undergo a GCT or OGTT. Conversely, a glucose loading test would have been avoided by 81.17% of women when FPG was set at 80 g/dL, but only 40.2% of GDM would have been detected.

Table 21: Screening for GDM using a fasting plasma glucose test

			Test accuracy							
Study	**Test**	**Threshold/cut-off**	**Sens. (95% CI)**	**Spec. (95% CI)**	**PPV (95% CI)**	**NPV (95% CI)**	**LR+**	**LR-**	**AUC (95% CI)**	**Ruling out GDM**
Maesa 2018[76]	FPG	55 mg/dL	97.8	1.3	1.39	97.62	0.99	1.69	0.633 (0.569 to 0.696)	84 (1.28)
	FPG	60 mg/dL	95.7	4.8	1.41	98.73	1.01	0.90		315 (4.79)
	FPG	62 mg/dL	91.3	10	1.42	98.79	1.01	0.87		659 (10.03)
	FPG	70 mg/dL	76.1	43.2	1.86	99.22	1.34	0.55		2,819 (42.89)
	FPG	80 mg/dL	40.2	81.5	2.99	98.97	2.17	0.73		5,335 (81.17)

Abbreviations: AUC: area under the curve; CI, confidence interval; FPG, fasting plasma glucose; GDM, gestational diabetes mellitus; LR: likelihood ratio; NPV: negative predictive value; PPV: positive predictive values

Maternal characteristics and risk factors

Four studies investigated the accuracy of testing for GDM using only maternal characteristics and risk factors and 1 study looked at the combination of maternal risk factors with the 1 h GCT (Table 22). The Farrar SLR included an analysis of the IPD of the BiB and ATLANTIC DIP cohorts, which is also presented in Table 22. Furthermore, they also included 24 studies screening by risk factors in their SLR chapter. However, due to high heterogeneity in the studies, Farrar 2016 did not conduct a MA, but provided a narrative summary of the included studies instead, which is also summarised below.[4]

Of the 24 studies in the Farrar 2016 SLR, 6 looked at the performance of existing riskbased guideline recommendations, 7 counted the number of risk factors each woman had, 6 used risk prediction models or scoring and 5 examined various risk factors. The outcome measure most common across the studies was the number of OGTTs required to diagnose a specific proportion of women with GDM. As expected, there appeared to be a linear correlation between the two; identifying more women with GDM requires offering more women an OGTT. Furthermore, the SLR found that no specific risk scoring was superior to another. It was recommended that using BMI and age would be the most

Furthermore, the authors remarked that the makeup of the population in terms of the risk factor prevalence will also affect the test accuracy, and populations where the risk factor prevalence is higher may benefit from universal, rather than risk-factor based screening, as most women would have been offered an OGTT test anyway. In populations with low risk factor prevalence, risk-factor based screening would be more effective to pick out women who should be offered the OGTT test.

The BEDIP-N study (n=1884) evaluated the accuracy of incorporating maternal risk into glucose testing using the GCT (threshold 7.2 mmol/L) for identifying GDM, including the risk factors of an ethnic minority background, a BMI ≥30 kg/m^2, a history of GDM, or any of these three risk factors.[80] In pregnant women who had any of these 3 risk factors in addition to a GCT value ≥7.2 mmol/L, the sensitivity for identifying GDM was 82.9% (95% CI 77.4 to 87.5) and the specificity was 57.5% (95% CI 55.0 to 59.9). While this combination achieved a higher sensitivity than using a single risk factor (ethnic minority background 78.1%, BMI 77.2%, history of GDM 74.1%) ot than screening with GCT only (72.4%), the specificity was compromised, resulting in a slightly lower positive likelihood ratio (LR+) of 1.9 (95% CI 1.8 to 2.1). By using a risk-factor based 2-step screening strategy, it was reported that the proportion of women requiring an OGTT based on a GCT would be reduced to 25.5%, with 52.6% of OGTTs potentially being avoided, compared with 1-step universal screening. Very similar sensitivity and specificity results were seen in Van Leeuwen 2010 (n=995) and Theriault 2014 (n=7208).[88, 89] Both studies used models, with Van Leeuwen 2010 building a predictive model on pre-pregnancy BMI, ethnicity, family history of diabetes and previous GDM, and Theriault 2014 validating 4 risk factor-based models (risk factors included age, BMI, ethnicity, history of diabetes/GDM/adverse obstetric outcomes). Van Leeuven 2010 found that their model performed better at identifying women with GDM and avoiding unnecessary OGTTs when the prevalence of GDM was assumed to be higher (≥4%) (number of OGTTs to diagnose 1 GDM was 11) but had a better sensitivity (75%) when the prevalence was lower (≥2%) (number of OGTTS for 1 GDM was 24). Theriault 2014 reported that the best performing model (by AUC) was Van Leeuwen's model; but while they found the specificity to be 80.7% (95% CI 79.6 to 81.8), the sensitivity was only at 60.4% (95% CI 54.3 to 66.1).

Saeedi 2018 (n=3616) looked at adding risk factors to FPG or even combining these with random blood glucose values. This approach had poor diagnostic power to detect GDM (highest sensitivity at 42% [95% CI 35 to 47]), though performed reasonably well at excluding women who did not have GDM from further testing (NPV at >90%).[79] By

contrast, the analysis of the IPD from BiB and ATLANTIC-DIP cohorts in Farrar 2016 showed a reasonably high sensitivity (95.9%) and PPV (84.5%) for age ≥25 years and BMI ≥25 kg/m^2, however, the sensitivity was low at only 16.5%.[4] Best specificity in Farrar 2016 was seen with the combination of age ≥25 years and BMI ≥25 kg/m^2 and previous GDM (24.6%) at which point the sensitivity was at 90.3% and the predictive value at only 76.5%. Of note, these results were still superior to screening by the risk factors recommended by the NICE guidelines, where sensitivity was at only 78.2% and specificity at 31.7%.

Temming 2016 (n=753) was the only study to investigate the combination of the GCT and maternal risk factors (history of GDM, age and BMI).[78] However, the combination of the two approaches did not improve the resulting test accuracy, as shown by the largest AUC reaching only 0.653 (95% CI 0.532 to 0.773).

Table 22. Screening for GDM using maternal risk factors

Study	Method of screening	Sens. (95% CI)	Spec. (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR– (95% CI)	AUC (95% CI)	Other
BEDIP-N[80]	Ethnic minority background + GCT (using a threshold of 7.2 mmol/L)	78.1 (72.1–83.3), 178/228	64.0 (61.6–66.3)	NR	NR	2.2 (2.0–2.4)	0.34 (0.27–0.44)	NR	Using any of the 3 risk factors, the proportion of women that would be missed would be reduced to 17.1% (n=39) 52.6% of all OGTTs could still be avoided
	BMI ≥30 kg/m^2 + GCT (using a threshold of 7.2 mmol/L)	77.2 (71.2–82.5)	64.3 (61.9–66.7)	NR	NR	2.2 (2.0–2.4)	0.35 (0.28–0.45)	NR	
	History of GDM + GCT (using a threshold of 7.2 mmol/L)	74.1 (67.9–79.7)	68.7 (66.4–71.0)	NR	NR	2.4 (2.1–2.6)	0.38 (0.30–0.47)	NR	
	Any of the 3 risk factors + GCT (using a threshold of 7.2 mmol/L)	82.9 (77.4–87.5)	57.5 (55.0–59.9)	NR	NR	1.9 (1.8–2.1)	0.30 (0.22–0.40)	NR	
Saeedi 2018[79]	Model I traditional risk factors*	28 (24 to 32)	86 (84 to 87)	20 (17 to 24)	90 (89 to 91)	NR	NR	0.43 (0.40 to 0.46)	NR
	Model I traditional risk factors* or RBG [random blood glucose] ≥8.0 mmol/L	36 (32 to 41)	84 (82 to 85)	23 (20 to 26)	91 (90 to 92)	NR	NR	0.40 (0.37 to 0/.43)	NR
	Model II traditional risk factors*	31 (25 to 37)	85 (84 to 86)	14 (11 to 17)	94 (93 to 95)	NR	NR	0.42 (0.38 to 0.46)	NR
Theriault 2014[88]	Naylor model: maternal age, prepregnancy BMI, ethnicity	72.2 (66.9 to 77.0)	55.1 (53.8 to 56.4)	8.2 (7.2 to 9.3)	97.3 (96.6 to 97.8)	NR	NR	0.668 (0.637 to 0.699)	NR
	Caliskan model: maternal age, pre-pregnancy BMI, prior adverse obstetric outcome, family history of diabetes, prior macrosomic fetus	71.1 (65.6 to 76.0)	59.3 (58.0 to 60.6)	9.3 (8.1 to 10.5)	97.2 (96.6 to 97.8)	NR	NR	0.680 (0.649 to 0.712)	NR
	Van Leeuwen model: prepregnancy BMI, ethnicity, family history of diabetes, previous GDM	60.4 (54.3 to 66.1)	80.7 (79.6 to 81.8)	14.9 (12.9 to 17.1)	97.3 (96.8 to 97.8)	3.13 (2.80–3.49)	0.49 (0.43–0.57)	0.756 (0.725 to 0.787)	NR
	Teede model: maternal age, BMI at first visit, ethnicity, family history of diabetes (1st degree), past history of GDM	65.6 (59.3 to 71.4)	75.0 (73.7 to 76.3)	13.5 (11.6 to 15.6)	97.3 (96.7 to 97.9)	NR	NR	0.739 (0.701 to 0.776)	NR
	Model II traditional risk factors* or RBG ≥8.0 mmol/L	41 (35 to 47)	83 (82 to 84)	16 (13 to 19)	95 (94 to 96)	NR	NR	0.38 (0.34 to 0.42)	NR

Temming 2016[78]	NDDG ≥160 mg/dL + history of GDM, age ≥30, BMI ≥30 kg/m^2	65.3% (50.4 to 78.3)	65.2% (42.7 to 83.6)	80.0% (64.4 to 90.9)	NR	NR	NR	0.653 (0.532 to 0.773)	NR
	NDDG ≥180 mg/dL + history of GDM, age ≥30, BMI ≥30 kg/m^2	38.8% (25.2 to 53.8)	82.6% (71.8 to 90.3)	82.6% (61.2 to 95.0)	NR	NR	NR	0.607 (0.502 to 0.712)	NR
	CC ≥160 mg/dL+ history of GDM, age ≥30, BMI ≥30 kg/m^2	56.9% (44.0 to 69.2)	57.1% (18.4 to 90.1)	92.5% (79.6 to 98.4)	NR	NR	NR	0.570 (0.363 to 0.777)	NR
	CC ≥180 mg/dL+ history of GDM, age ≥30, BMI ≥30 kg/m^2	32.3% (21.2 to 45.1)	71.4% (29.0 to 96.3)	91.3% (72.0 to 98.9)	NR	NR	NR	0.519 (0.329 to 0.708)	NR
Van Leeuwen 2010[89]	Universal testing (age, BMI, non-Caucasian ethnicity, smoking, previous miscarriage	100	100	NR	NR	NR	NR	NR	NND (N/n omen with GDM): 42 (995/24)

	history of: diabetes or GDM or perinatal death).								
	Diagnostic testing if probability of GDM ≥2% (risk factors as above)	75.0 (55.4 to 88.0)	57.8 (57.3 to 58.1)	4.2 (3.1 to 4.9)	98.9 (98.1 to 99.5)	1.78 (1.30 to 2.10)	0.43 (0.21 to 0.78)	NR	NND (N/n women with 24 (428/18)
	Diagnostic testing if probability of GDM ≥4% (risk factors as above)	45.8 (28.2 to 64.5)	88.4 (87.9 to 88.8)	8.9 (5.5 to 12.5)	98.5 (98.0 to 99.0)	3.94 (2.34 to 5.77)	0.61 (0.40 to 0.81)	NR	NND (N/n women with GDM): GDM): 11 (124/11)
Farrar 2016 (BiB and Atlantic DIP cohorts)[4]**	Age ≥30 years, BMI ≥25 kg/m^2, diabetes, prior GDM	90.3	24.6	76.5	NR	NR	NR	NR	NR
	Age ≥25 years, BMI ≥25 kg/m^2	95.9	16.5	84.5	NR	NR	NR	NR	NR
	Age ≥25 years, BMI ≥25 kg/m^2, prior GDM	95.9	16.5	84.5	NR	NR	NR	NR	NR
	NICE guideline recommended risk factors	78.2	31.7	67.2	NR	NR	NR	NR	NR

*Model I (modified IADPSG criteria), 1.75 OR of adverse events in HAPO: equivalent cFPG ≥4.6 mmol/L or 2h OGTT ≥8.5 mmol/L. Model II, 2.0 OR of adverse events in HAPO: equivalent cFPG ≥4.8 mmol/L or 2h OGTT ≥ 9.0 mmol/L. Traditional risk factors = heredity (first-degree relative with diabetes), obesity (pre-pregnancy weight ≥90 kg), previous LGA infant (≥4500 g or ≥mean + 2SD), previous GDM. **Only results with the highest sensitivity or specificity are presented; for full results please refer to the data extraction tables in Appendix 3

Abbreviations: AUC: area under the curve; BMI, body mass index; CC, Carpenter and Coustan; CI, confidence interval; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; LR: likelihood ratio; OGTT, oral glucose tolerance test; NDDG, National Diabetes Data Group; NICE, National Institute for Health and Care Excellence; NND, number needed to diagnose; NPV: negative predictive value; NR, not reported; PPV: positive predictive values; RBG, random blood glucose

Biomarkers

Three studies examined the feasibility of screening for GDM using blood-based biomarkers other than glucose. Khalafallah 2016 (n=480) explored the accuracy of HbA1c for prediction of GDM at various cut-off values, finding that a cut-off of 5.4% achieved a NPV of 91% and specificity of 95%, although with a very low sensitivity of 26.5%.[82] Fructosamine demonstrated poor predictive value in the Gingras 2018 study (n=1488), with an AUC of 0.52.[83] The highest sensitivity for detection of GDM was achieved at a cut-off of ≥222 µmol/L (≥50th percentile) with 48.6% specificity. Using a cut-off at the 75th and 95th percentiles in order to increase specificity (74.9% and 95.1% respectively) substantially decreased sensitivity. These findings suggest poor suitability of fructosamine as a screening test for GDM, due to poor sensitivity to detect abnormal glucose tolerance. Similarly, AUC values appeared similar across all lipid and apolipoprotein markers investigated by Iimura 2015 (n=266), with levels of triglycerides achieving the highest accuracy (0.624, 95% CI 0.490 to 0.759).[84] The authors concluded that none of the markers demonstrated sufficient accuracy for prediction of GDM.

Table 23. Screening for GDM using biomarkers

Study	Test/Biomarker	Threshold	Sens. (95% CI)	Spec. (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
Khalafallah 2016[82]	HbA1c	10%	0	99.7	0	88.8	NR
		6.1%	2	99.7	50	89	NR
		6%	4.1	99.7	66.7	89.2	NR
		5.9%	6.1	99.7	75	89.4	NR
		5.8%	8.2	99.7	80	89.6	NR
		5.7%	10.2	99.5	71.4	89.8	NR
		5.6%	12.2	99	60	90	NR
		5.5%	22.4	98.2	61.1	91	NR
		5.4%	26.5	95.4	41.9	91.2	NR
		5.3%	34.7	88.4	27.4	91.5	NR
		5.2%	55.1	79.7	25.5	93.4	NR
		5.1%	61.2	67.6	19.2	93.3	NR
		5%	69.4	51.9	15.4	93.1	NR
		4.9%	73.5	31.4	11.9	90.4	NR
		4.8%	81.6	18	11.1	88.6	NR
		4.7%	95.9	10	11.8	95.1	NR
		4.6%	95.9	4.6	11.2	90	NR
Gringras 2018 (Project Viva)[83]	Fructosamine	≥222 µmol/L (≥50th percentile)	54.8	48.6	5.2	95.4	0.52
		≥256 µmol/L (≥75th percentile)	26.0	74.9	5.1	95.2	NR
		≥312 µmol/L (≥95th percentile)	6.9	95.1	6.7	95.2	NR
Iimura 2015[84]	Lipid and apolipoprotein markers	TG	NR	NR	NR	NR	**0.624 (0.490 to 0.759)**
		ApoC-III	NR	NR	NR	NR	0.583 (0.451 to 0.715)
		ApoB48	NR	NR	NR	NR	0.568 (0.439 to 0.697)
		ApoA-I	NR	NR	NR	NR	0.560 (0.438 to 0.684)

		HDL-C	NR	NR	NR	NR	0.531 (0.420 to 0.641)
		ApoB	NR	NR	NR	NR	0.519 (0.391 to 0.648)
		TC	NR	NR	NR	NR	0.518 (0.388 to 0.647)
		LDL-C	NR	NR	NR	NR	0.515 (0.393 to 0.636)

Abbreviations: Apo, apolipoprotein; AUC: area under the curve; CI, confidence interval; HDL-C, high-density lipoprotein C; LR: likelihood ratio; NPV: negative predictive value; NR, not reported; PPV: positive predictive values; TC, total cholesterol; TG, triglyceride

Conclusions

Screening by risk factors only did not appear to be a valid strategy to detect GDM, either when combined with FPG values, as demonstrated by Saeedi 2018 or with GCT, as shown in Temming 2016.[78, 79] In fact, the best performance with an AUC of 0.911 (95% CI 0.868 to 0.954) was achieved by the 2 h 100 g OGTT following a GCT, which is currently used in clinical practice and recommended by some of the guidelines.[26, 36] The problem is that the reference standard was the same test, only using a pre-specified threshold, so it appears that currently the accuracy hinges more on the thresholds used, than the test. Without another test or clinical diagnosis of GDM, it is not possible to reliably ascertain the validity of the OGTT test. Furthermore, this test involved a GCT, which includes an additional glucose loading step, which is problematic due to the potential side-effect of glucose loading especially in women who may have elevated glucose and lower glucose tolerance.
Use of only OGTT as a screening test was only reported by 1 study (Ryser Ruetschi 2016),[81] where it did not appear to perform differently to a risk factor based approach combined with a GCT (78.5% sensitivity/69% specificity with OGTT only vs e.g. 82.% sensitivity/57.5% specificity with 3 risk factors + GCT in BEDIP-N[80]). Other risk factor based tests were also similar in terms of sensitivity and specificity combinations, but comparisons are difficult as the Ryser Ruetschi 2016[81] study did not report an AUC.

Furthermore, the study by Kosus 2012 was conducted in Turkey and with the aim to understand whether a better diagnostic threshold could be used in the diagnosis of GDM using the OGTT within that specific population. Importantly, several studies have remarked that the most fitting screening strategy may be dependent on the prevalence of GDM and maternal risk factors within a population. For example, in populations with a higher GDM prevalence, universal screening may be more effective than in low-GDM populations, where a risk-factor based approach could be used. Therefore, the performance of the tests may be improved depending on whether they are used in universal or targeted screening, and thus studies of test accuracyneed to be interpreted by considering both the prevalence of GDM in the study cohort and the prevalence of GDM in the population to which the test would be applied. In this rapid review for example, only 1 study was conducted in the UK, and whilst the prevalence of GDM in the countries where other studies were performed was assumed to be similar to that of the UK, whether small differences can affect test performance

remains uncertain. Future evidence syntheses may be more relevant if they only include studies on the unselected UK pregnant women.

Outside of the population applicability issues, the main issue appears to be that for any test, the higher thresholds can increase specificity, avoiding the OGTT in many women, but also missing a significant proportion of women with GDM. On the other hand, low thresholds can achieve a high sensitivity, but their specificity is low, leading to unnecessary OGTTs, which could adversely affect some women, for example those who may be glucose intolerant, but not have GDM. Furthermore, the test involves repeated blood draws and requires overnight fasting, which may not be acceptable to many women.

Summary of Findings Relevant to Criterion 4: Criterion not met†

Quantity: The evidence base consisted of the Farrar 2016 IPD analysis and SLR, as well as of 13 primary studies. The only UK cohort was included in the IPD (BiB), otherwise 2 studies were from the US, 2 from Japan and the remaining 9 primary studies in a different country each. The number of women screened varied between 202 and 7,208 in the primary studies;[88] 16,537 women were also included in the IPD analysis of Farrar 2016. Studies focused mostly on the accuracy of using various maternal characteristics or risk factors in screening for GDM (5 studies and Farrar 2016), or on the accuracy of glucose tests (5 studies). Three studies investigated use of biomarkers in GDM detection. None of the tests had performance superior to that of the currently used reference standard and diagnostic test (GCT followed by 2 h OGTT using 151.5 mg/dL as the threshold for elevated glucose). Screening by maternal risk factors, GCT or FPG was less accurate with performance somewhat variable and depending on the risk factors and glucose thresholds used, whereas screening using biomarkers had the poorest performance.

Quality: Studies were generally of good quality and at a low risk of bias. Issues in the primary studies were mostly around the index test and not reporting a pre-specified threshold as well as not offering a reference standard to all women in the study. Screennegative women were mostly not offered the reference standard resulting in partial verification, possibly leading on the overestimation of the performance in the diagnostic test accuracy. The Farrar SLR was judged to be of high quality and at a low risk of bias.

† **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.

Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

Applicability: Although only the BiB cohort in Farrar 2016 was from the UK, the the evidence was generally considered applicable to the UK setting. There were 5 studies with some concerns around applicability due to the included population potentially having a different ethnic make-up and in two studies the tests used (both biomarkers) were not ones commonly used in UK clinical practice for population screening.

Consistency: Five studies investigated maternal risk factors and five looked at glucose tests; however, consistency among studies of maternal risk factors was difficult to assess as the strategies differed both with respect to the risk factors and the glucose tests and thresholds used. Nevertheless, the prevailing results were consistent, in that the higher thresholds produced a high specificity, but low sensitivity and lower thresholds had the opposite effect on the test parameters. However, some GCT results had better specificity with a lower threshold than others, though this may have been due to the use of different reference standard criteria or just heterogeneity between the studies. None of the 3 studies of biomarkers evaluated the same test, thus consistency could not be determined.

Summary: The glucose loading OGTT test was found to have a superior performance to any other test; of the studies found a screening strategy that achieved test accuracies where both specificity and sensitivity were high enough to consider the test as reliable and able to replace the current test (2 h 75 g OGTT), which involves glucose loading and therefore poses some risk of harm to women who are already suspected to be at risk of glucose intolerance. Using any of those strategies and only applying OGTT in screenpositive women would likely miss a considerable proportion of GDM (at a high threshold) or result in most women having to undergo OGTT anyway (at a lower threshold). Therefore, the best currently available test is the diagnostic OGTT test. Its drawbacks are uncertainty around its accuracy versus a different reference standard or clinical diagnosis, as well as the risk of harm, with unknown consequences should it be used in the population of all pregnant women. Given the uncertainty around the accuracy of the OGTT, its unclear acceptability if used for screening, and lack of any better screening test, criterion 4 is not met.

Criterion 9 — There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care.

Question 3 – What is the most effective intervention for lowering glucose levels in screen-detected pregnant women with GDM and preventing adverse perinatal outcomes?

The aim of question 3 was to identify the efficacy of interventions – compared with other interventions, no treatment or usual care – for lowering glucose levels and preventing adverse outcomes in pregnant women with screen-detected GDM (i.e. low-risk women who would not currently be treated based on the current NICE guidance).

The rapid review conducted for the UK NSC by Waugh and colleagues in 2010 synthesised evidence for primary studies comparing insulin with usual care and oral glucose-lowering treatment (with antidiabetic agents) with insulin for hyperglycaemia in pregnancy. It was not necessary for GDM to be screen-detected and, in the majority of studies, it was not specified whether screening was used. Two randomised trials compared insulin with usual care. ACHOIS was an Australia- and UK-based trial where women with diagnosed gestational diabetes were randomised to receive dietary advice and advised to self-monitor glucose (the intervention group) or to standard care. In the intervention group, 20% of women commenced insulin therapy. The rate of serious perinatal complications including death, shoulder dystocia, bone fracture and nerve palsy was significantly lower in the intervention group compared with the control group. A second study, the MFMU Network trial compared diet and insulin therapy (if required) with "no specific treatment" in women with screen-detected mild hyperglycaemia in pregnancy. In contrast with ACHOIS, there was no significant difference between the intervention and control groups for the primary outcome of a composite of perinatal mortality and morbidities (stillbirth, neonatal mortality, hypoglycaemia, hyperbilirubinaemia, neonatal hyperinsulinaemia and birth trauma). Waugh *et al.* noted that this difference may have been due to women in the MFMU Network trial having lower levels of hyperglycaemia than women in the ACHOIS trial.[5]

For oral antidiabetic agents compared with insulin, 7 RCTs and 20 cohort studies were included, comprising a total of 4425 participants. A large range of maternal and neonatal outcomes were reported. Overall, the RCT evidence showed little difference between oral drugs and insulin. In a comparison between glibenclamide and insulin, maternal hypoglycaemia was lower, whereas neonatal hypoglycaemia and birth weight was lower for glibenclamide. In a comparison between metformin and insulin, maternal weight gain was lower with metformin, but age at delivery favoured insulin. Waugh 2020 concluded that both glibenclamide and metformin could be used as alternatives to insulin, but noted that the evidence base at the time was not sufficient to enable decision making about when the best time to initiate therapy was.[5]

Eligibility for inclusion in the review

This review searched for interventional and observational studies, including prospective, retrospective and case-control studies. SLRs and MAs of these relevant study types, published since January 2009, were also eligible for inclusion. Studies were included if the population comprised pregnant women with GDM, or newborns of women with GDM (fewer than 28 days of age). In order to avoid limiting available evidence, studies on any pregnant women with GDM, not only screen-detected GDM were included. However, populations with screen-detected treatments of interest included both pharmacological and lifestyle interventions for GDM (e.g. diet and/or exercise). Comparators could be another treatment, placebo or no treatment/standard of care. Outcomes of interest for question 3 included, but were not limited to, pregnancy outcomes such as perinatal mortality, mode of birth and gestational weight gain; maternal outcomes such as postpartum haemorrhage, method of infant feeding and post-pregnancy type 2 diabetes; and neonatal outcomes including macrosomia, birth injury, hypoglycaemia and admission to the neonatal intensive care unit (NICU). Studies were not restricted geographically. Full details of eligibility criteria are presented in Table 5.

Evidence was initially classified into 2 tiers: Tier 1 comprised studies conducted in the UK whereas Tier 2 included studies conducted in all other eligible (OECD or EEA) countries. Due to the high volume of evidence encountered, at the full text review stage the studies were also further classified into 2 tiers based on study design: Tier 1 included only RCTs (and SLRs/MAs thereof), and Tier 2 contained any other study design (and SLRs/MAs thereof).

For this question (Q3), 3 SLRs on anti-diabetics, insulin or lifestyle interventions for GDM were included as an evidence base and updated, with searches date-limited to 2016, when the searches for these 3 SLRs were run.[91-93] In addition, chapter 6 of Farrar 2016, an SLR and MA that formed part of the evidence base for Q1 and Q2 of this review (with searches run in 2014), was also included *a priori*.[4] Any RCTs captured in these SLRs were not dataextracted separately; their results were only included as part of any pooled/MA conducted in the SLR, to avoid duplicate inclusions of the same trial.

Description of the evidence

Characteristics of included studies (Q3)

A total of 4 SLRs and 59 publications on 55 studies were included. The SLRs included 26 unique RCTs.

Due to the high volume of evidence identified, Tier 1 study design records, which comprised RCTs from any eligible country (12 publications on 8 unique RCTs), were prioritised for evidence synthesis. Prioritising only RCT evidence was also in line with the approaches taken in the 4 included SLRs. Due to lower methodological quality, the non-randomised interventional and observational studies (Tier 2, 47 publications) whilst included in the review, were not extracted or considered in the evidence synthesis.

None of the SLRs specified that women with GDM were required to be screen-detected; women just needed to have GDM, with diagnoses as defined by individual studies. All but 1 of the 8 included RCTs also did not specify whether the population needed to be screendetected GDM. The only study that did was the MFMU Network RCT, which noted that women without an overt diagnosis of diabetes mellitus underwent universal screening with a 50 g GCT between 24 and 28 weeks of gestation.[94]

Brown 2017L/A/I and Farrar 2016 SLRs

Three high quality SLRs conducted by the Cochrane Collaboration were part of the evidence base for this question. The aim of Brown 2017L (Lifestyle) was to evaluate the effect of lifestyle interventions with or without pharmacotherapy on treatment of women with GDM.[91] Brown 2017A (Anti-diabetics)[93] evaluated the effect of anti-diabetic pharmacological therapies and Brown 2017I (Insulin)[92] looked specifically at the effect of insulin. Farrar 2016 examined the effect of all of the above on the treatment of GDM (lifestyle interventions, insulin and other pharmacological therapies).[4]

The scopes of these reviews were very closely aligned with the eligibility criteria for question 3, except that the SLRs were broader in that they had no geographical limits on where the studies were conducted, and examined more outcomes than were the focus for this rapid review.

Brown 2017L included 15 RCTs of lifestyle interventions in women with GDM, conducted in the US (n=4), China (n=2), Iran (n=2), Canada (n=2), UK (n=1), Italy (n=1), UAE (n=1), Thailand (n=1), and Australia and UK (n=1).[91] Nine of the included RCTs provided details of diagnostic criteria for GDM, which included WHO 1999 (n=3); CC (n=2); ADA 2000 (n=1); ADIPS 1998 (n=1); IADPSG 2010 (n=1) and Hatem 1988 (n=1). Six RCTs did not provide details of diagnostic criteria.

Brown 2017A included 11 RCTs in the qualitative synthesis and 8 RCTs in the MA (quantitative synthesis) on antidiabetic agents (metformin, glibenclamide, acarbose, chlorpropamide and tolbutamide) in women with GDM.[93] The RCTs were conducted in Brazil (n=3), US (n=3), India (n=2), South Africa (n=1), UK (n=1) and Israel (n=1). The

diagnostic criteria for GDM used in individual studies was reported in 8 RCTs, and comprised CC (n=3); NDGG 1979 (n=2); WHO 1999 (n=2) and unspecified WHO (n=1). Three RCTs did not provide details of diagnostic criteria.

Brown 2017I included 53 RCTs in the qualitative synthesis and 51 RCTs in the MA (quantitative synthesis) with at least 1 insulin therapy arm.[92] The RCTs were conducted in the US (n=16), India (n=7), Iran (n=6), Egypt (n=3), Brazil (n=3), Pakistan (n=3), Finland (n=3), Italy (n=2), Sweden (n=1), Canada (n=1), Ghana (n=1), Australia (n=1), New Zealand and Australia (n=1), Turkey (n=1), Israel (n=1), Malaysia (n=1), South Africa (n=1) and Poland (n=1). The diagnostic criteria for GDM were reported in 35 studies, and not reported in 18 studies.

The Farrar 2016 SLR/MA itself updated 5 existing SLRs and included 47 RCTs in the qualitative synthesis and 45 RCTs in the MA (quantitative synthesis) and included any of lifestyle, insulin or antidiabetic interventions.[4] Twenty-three trials included antidiabetic agents in at least 1 arm (metformin and/or glibenclamide), 5 trials compared different insulin formulations, 9 trials compared different diets and 10 trials compared combinations of diet modification, glucose monitoring and insulin with routine obstetric care. The diagnostic criteria for GDM was varied and included CC, NDDA, WHO, ADA or local guidelines.

Primary RCTs from database searches

Ultimately, 12 articles on 8 unique RCTs from databases were selected for extraction for question 3.[94-105] The smallest study recruited 12 participants, and the largest study recruited 932 participants.[94, 95] Identified evidence was found for the following treatments: insulin, metformin, glibenclamide, glyburide, dietary interventions including low or high carbohydrate diets, and a structured exercise programme.

Four RCTs reported on the impact of pharmacological treatments on pregnancy and neonatal outcomes in women with GDM.[4, 91, 92, 96, 97, 99, 100] Each of the RCTs was conducted in a different country: GRACES in the UK, INDAO in France, Pellonpera 2016 in Finland, and MiG in Australia and New Zealand.[96, 97, 99, 100] The biggest sample was enrolled in INDAO (N=809) and the smallest in GRACES (N=23).[96, 97] Two trials compared glyburide with insulin (INDAO and GRACES) and 2 compared metformin with insulin (Pellonpera 2016 and MiG).[99, 100]

Four RCTs evaluated the impact of lifestyle interventions on pregnancy and neonatal outcomes in pregnant women with GDM.[91, 94, 95, 98, 104] One RCT, Kokic 2018, was conducted in Croatia and compared a structured exercise programme plus nutritional therapy against standard prenatal care.[98] The other 3 RCTs were all conducted in the US

and investigated the effects of diet or nutritional advice on GDM treatment.[94, 95, 104] In the MFMU Network trial, women were either randomised to formal nutritional counselling and diet therapy along with insulin if required, or usual prenatal care.[94]

In the Trout 2016 study, women with GDM randomised to the intervention group were instructed on minimum and maximum recommended carbohydrate levels (35 to 40% of total calories, respectively). Women in the control group had a carbohydrate intake level set at 50–55% of total calories.[104] The CHOICE diet study compared a higher-complex carbohydrate, lower-fat diet (CHOICE diet, composed of 60% carbohydrate, 25% fat, 15% protein) with a low-carbohydrate, higher-fat diet (composed of 40% carbohydrate, 45% fat, 15% protein, matched with the CHOICE diet for fat, simple sugars and fibre content). Menus were prepared by the research centre nutrition serviced and picked up by participants every 72 hours.[95]

A study-level summary of data extracted from each included publication is presented in Appendix 3.

Discussion of findings

Quality assessment

Brown 2017 and Farrar 2016 SLRs

The quality of the 4 included SLRs was appraised using the AMSTAR 2 checklist; a summary is presented in Table 102 (Appendix 4).

The SLRs were found to be at a low risk of bias. All sufficiently described the objectives and inclusion criteria using the PICO framework, and had their methods established prior to commencing the review as evidenced by the availability of protocols. While all 4 SLRs were failed to provide a justification of study design selected, this is unlikely to place them at a high risk of bias as their selection of studies was appropriate for the research question posed. Search strategy, study selection and data extraction were judged to be appropriate and reporting was comprehensive in all SLRs except for Farrar 2016, who did not report on the source of funding for included studies. All SLRs also conducted MAs using appropriate statistical methods, and sufficiently assessed the risk of bias of the individual studies and the potential impact on results. Although Brown 2017A and Brown 2017L failed to investigate publication bias, it is not expected to affect the applicability of the reviews.

RCTs

The quality of the 8 included RCTs (reported through 12 publications) was appraised using an adapted Cochrane Risk of Bias checklist,[106] (Table 96; Appendix 4). A summary of the risk of bias is presented in Table 24, and the full appraisal is presented in Table 106 to Table 108 (Appendix 4). Overall, 3 studies were judged to be at low risk of bias, 3 had some concerns of bias and 2 were at high risk of bias due to issues with missing outcome data in 1 case and measurement of outcome in the other.

Table 24. Summary of Cochrane Risk of Bias assessments for RCTs evaluating treatment in women with GDM

	Pharmacological interventions					Lifestyle interventions		
	Rowan 2018 [99] (MiG)	**Hernande z 2016 (CHOICE [95, 101])**	**2017 Kokic (INDAO)**	**Senat Pellonpera [97]**	**Reynolds [94, 96] (GRACES)**	**2016 Risk of bias 2018**	**2018 [98, 100, 103]**	**Trout [104] 2016 Network [102] MFMU diet)**
Randomisation process	Low	Some concerns	Low	Low	Low	Some concerns	Some concerns	Some concerns
Effect of assignment to intervention	Low	Low	Low	Low	Low	Low	Low	Some concerns
Missing outcome data	High	Low	Low	Low	Low	Low	Low	Low
Measurement of outcome	Low	High	Low	Low	Low	Low	Low	Low
Selection of the reported result	Some concerns	Low	Low	Low	Low	Some concerns	Some concerns	Some concerns
Overall risk of bias	**High risk of bias**	**High risk of bias**	**Low risk of bias**	**Low risk of bias**	**Low risk of bias**	**Some concerns**	**Some concerns**	**Some concerns**

Randomisation process

The risk of bias arising from the randomisation process was judged to be low across 4 trials, and at "some concerns" for the other 4 trials, where reporting of the randomisation and allocation concealment was poor.[94, 95, 100] However, randomisation was deemed appropriate as demonstrated by similar baseline characteristics between treatment arms.

Effect of assignment to interventions

There was a low risk of bias in 7 out of 8 included trials for this domain. None of the studies were reported to have been blinded, and therefore study personnel and participants were likely aware of treatment allocation. However, this was not judged to adversely impact

assignment to interventions as there were no apparent deviations from the intended intervention due to the lack of blinding. All trials analysed outcome data on an intentiontotreat (ITT) or modified ITT basis, demonstrating appropriate analysis methods. There was some concern about the risk of bias due to deviations form interventions in Trout 2016, as a considerable proportion of women did not complete their food logs but there was no information on likely deviations from diet in these participants.[104]

Missing outcome data

Outcome data was available for at least 90% of participants in 5 out of the 8 included trials. In the MiG trial, the long-term follow-up rate was low and considered to be different from the initial cohort.[99] This study was therefore at a high risk of bias for this domain.

Measurement of outcome

The CHOICE trial was at a high risk of bias for this domain, as the analysis of study outcomes was not sufficiently powered for statistical analyses due to a small sample size.[95] The methods of measuring outcomes were considered appropriate and consistent between treatment arms in the other 7 included trials, resulting in low risk of bias.

Selection of the reported result

Three trials carried some concerns for bias for selection of the reported result, due to unavailability of a pre-specified analysis plan.[94, 99, 100] The other 5 trials reported that outcomes were pre-specified and were therefore at a low risk of bias.

Results

Key results are presented in Table 41. Full details of the included studies and their results can be found in Appendix 3. In the following sections, outcome results are considered separately for the different comparisons of interventions of interest.

Glibenclamide vs placebo

One RCT included in the Brown 2017A SLR compared the effects of treatment with glibenclamide vs placebo. This was the only study identified in the rapid review that examined the comparison between an oral antidiabetic agent and placebo. There was no significant difference between glibenclamide and placebo for any of the reported maternal/pregnancy or neonatal outcomes, with all 95% CIs of RR spanning from below to above 1 (Table 25 and Table 26).[93]

Table 25. Maternal and pregnancy outcomes reported by trials of glibenclamide vs placebo for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Pregnancy outcomes						
Preeclampsia	Brown 2017A (1 RCT)[93]	Placebo	Anticipated absolute effects	167 per 1000	1 (ref)	NR
		Glibenclamide		207 per 1000 (95% CI 135 to 317)	1.24 (0.81 to 1.90)	
Mode of delivery						
Induction of labour	Brown 2017A (1 RCT)[93]	Placebo	Anticipated absolute effects	188 per 1000	1 (ref)	NR
		Glibenclamide		222 per 1000 (95% CI 149 to 331)	1.18 (0.79 to 1.76)	
C-section		Placebo		360 per 1000	1 (ref)	NR
	Brown 2017A (1 RCT)[93]	Glibenclamide	Anticipated absolute effects	371 per 1000 (95% CI 285 to 483)	1.03 (0.79 to 1.34)	

Abbreviations: CI, confidence interval; NR, not reported; RCT, randomised controlled trial; RR, relative risk

Table 26. Neonatal outcomes reported by trials of glibenclamide vs placebo for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Glucose tolerance						
LGA	Brown 2017A (1 RCT)[93]	Placebo	Anticipated absolute effects	118 per 1000	1 (ref)	NR
		Glibenclamide		105 per 1000 (95% CI 60 to 187)	0.89 (0.51 to 1.58)	
Neonatal hypoglycaemia	Brown 2017A (4 RCTs)[93]	Glibenclamide	Anticipated absolute effects	11 per 1000	1 (ref)	NR
		Metformin		21 per 1000 (95% CI 4 to 114)	1.97 (0.36 to 10.62)	

Abbreviations: CI, confidence interval; LGA, large for gestational age; NICU, neonatal intensive care unit NR, not reported; RCT, randomised controlled trial; RR, relative risk

Metformin vs insulin

Two RCTs and 1 SLR (Farrar 2016) compared metformin and insulin.[4, 99, 100]

Maternal and pregnancy outcomes

Two RCTs (the Australian 2018 MiG study[99] and Finnish Pellonpera 2016[100]) and the Farrar 2016 SLR[4] (including data on 3 to 5 RCTs, depending on the outcome) compared metformin and insulin and reported on at least 1 maternal or pregnancy outcome (Table 27).

There was no significant difference between metformin and insulin in preventing gestational hypertension in either of the 2 trials that reported this outcome. The MiG study found rates of 1% vs 0% (p=1.00) in the 7-year cohort, and 11.1% vs 5.5% (p=0.46) in the 9-year cohort for metformin and insulin, respectively,[99] whilst Pellonpera 2016 found rates of 1.8% vs 3.7% (p=0.44).[100] The same 2 RCTs also reported on rates of pre-eclampsia and similarly, found no significant difference between the interventions (metformin vs insulin, 5.1% vs 3.9% [p=1.00] in MiG 7-year cohort; 4.4% vs 0% [p=0.20] in MiG 9-year cohort; 4.6% vs 9.3% [p=0.17]). These findings were reflected by the Farrar 2016 meta-analysis which found no clear difference in risk of pre-eclampsia in women treated with metformin compared to women treated with insulin (RR 0.74, 95% CI 0.48 to 1.14).[4] The MiG study and Pellonpera 2016 also both reported on gestational age at birth, which ranged from 38.4 ± 1.3 to 39.2 ± 1.4 weeks in women treated with metformin,[99, 100] and 38.5 ± 1.2 to 39.4 ± 1.6 weeks in women treated with insulin.[99, 100] Comparisons were insignificant with the exception of the 7-year MiG cohort where gestational age was slightly higher in women treated with insulin than with metformin (38.8 ± 1.0 weeks vs 38.4 ± 1.2 weeks, p=0.05).[99] In similar fashion, there were no significant differences in rates of pre-term birth in either the MiG study (metformin vs insulin, 10.3% vs 3.9% [p=0.28] in the 7-year cohort; 11.1% vs 11.1% [p=1.00] in the 9-year cohort)[99] or the Farrar 2016 SLR (RR 1.37, 95% CI 0.62 to 3.01 for metformin vs insulin).[4]

The Farrar 2016 MA and Pellonpera 2016 reported comparisons between metformin and insulin for 3 different modes of birth: assisted/instrumental vaginal delivery; induction of labour and C-section.[4, 100] The results from both studies were in agreement for C-section, in that there was no significant difference between the rates of these outcomes in women treated with metformin or insulin (p>0.05) (Table 27). However, results were inconsistent for instrumental delivery and induction of labour. The Farrar 2016 MA found that the risk of instrumental delivery was significantly higher in women treated with metformin than insulin (RR 1.66, 95% CI 1.37 to 2.01) based on data from 3 trials,[4] whilst Pellonpera 2016 found no significant difference (metformin vs insulin, 8.3% vs 7.5%, p=0.83).[100] This pattern was reversed for induction of labour. Whilst Pellonpera 2016 found that induction of labour was significantly more common in women treated with insulin than metformin (54.2% vs 37.6%, p=0.014),[100] Farrar 2016 saw no such association (RR 0.84, 95% CI 0.60 to 1.18 for metformin vs insulin).[4]

The MiG 2018 study and Pellonpera 2016 also reported on methods of infant feeding. Neither measures of risk nor levels of significance were reported for breastfeeding outcomes, therefore differences between insulin and metformin treatment could not be quantified. However, in the MiG 2018 study, rates of breastfeeding and formula feeding were similar in both arms in the 7- and 9-year cohorts (Table 27).[99] In Pellonpera 2016, the mean duration of breastfeeding following delivery was also similar between treatment arms for breastfeeding overall (metformin: 6.31 ± 4.00 months vs insulin: 6.59 ± 4.44 months),

and for exclusive breastfeeding (metformin: 2.76 ± 2.37 months vs insulin: 2.58 ± 2.43 months).[100]

The only long-term maternal outcome reported for the comparison between metformin and insulin was post-pregnancy type 2 diabetes in Pellonpera 2016. The rates appeared similar (3.9% vs 5.0%) but level of significance was not reported.[100]

Metformin and insulin appear comparable in terms of maternal and pregnancy outcomes. Nevertheless, as no studies were conducted specifically in screen detected women, it is uncertain whether the same conclusion could be drawn for this population.

Page

Table 27. Maternal and pregnancy outcomes reported by trials of metformin vs insulin for GDM

Outcome	Study	Intervention/Comparator	Outcome measure	Outcome value	RR (95% CI)	p-value
Pregnancy outcomes						
Gestational hypertension	Pellonpera 2016[100]	Metformin (n=110)	n (%)	2 (1.8)	NR	0.44
		Insulin (n=107)		4 (3.7)	NR	
	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109)	n (%)	1 (1.7)	NR	1.00
		Insulin, 7-year cohort (n=51)		0 (0)	NR	
		Metformin, 9-year cohort (n=45)		5 (11.1)	NR	0.46
		Insulin, 9-year cohort (n=54)		3 (5.5)	NR	
Pre-eclampsia	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109)	n (%)	3 (5.1)	NR	1.00
		Insulin, 7-year cohort (n=51)		2 (3.9)	NR	
		Metformin, 9-year cohort (n=45)		2 (4.4)	NR	0.20
		Insulin, 9-year cohort (n=54)		0 (0)	NR	
	Pellonpera 2016[100]	Metformin (n=110)	n (%)	5 (4.6)	NR	0.17
		Insulin (n=107)		10 (9.3)	NR	
	Farrar 2016; 4 RCTs	Metformin	NR	NR	0.74 (0.48 to 1.14)	NR
		Insulin	NR	NR	1 (ref)	
Gestational age at birth	Pellonpera 2016[100]	Metformin (n=110)	Mean (SD)	39.2 (1.40)	NR	0.43
		Insulin (n=107)		39.4 (1.58)	NR	
	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109)	Mean (SD)	**38.4 (1.2)**	NR	**0.05**
		Insulin, 7-year cohort (n=51)		**38.8 (1.0)**	NR	
		Metformin, 9-year cohort (n=45)		38.4 (1.3)	NR	0.75
		Insulin, 9-year cohort (n=54)		38.5 (1.2)	NR	
Pre-term birth	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109)	n (%)	6 (10.3)	NR	0.28
		Insulin, 7-year cohort (n=51)		2 (3.9)	NR	
		Metformin, 9-year cohort (n=45)		5 (11.1)	NR	1.00
		Insulin, 9-year cohort (n=54)		6 (11.1)	NR	
	Farrar 2016; 4 RCTs	Metformin	NR	NR	1.37 (0.62 to 3.01)	NR
		Insulin	NR	NR	1 (ref)	

Mode of birth						
Assisted/instrumental vaginal	Pellonpera 2016[100]	Metformin (n=110)	n (%)	9 (8.3)	NR	0.83
		Insulin (n=107)		8 (7.5)	NR	
	Farrar 2016; 3 RCTs	Metformin	NR	NR	**1.66 (1.37 to 2.01)**	NR
		Insulin	NR	NR	**1 (ref)**	
Induction of labour	Pellonpera 2016[100]	Metformin (n=110)	n (%)	**41 (37.6)**	NR	**0.014**
		Insulin (n=107)		**58 (54.2)**	NR	
	Farrar 2016	Metformin	NR	NR	0.84 (0.60 to 1.18)	NR
		Insulin	NR	NR	1 (ref)	
C-section	Pellonpera 2016[100]	Metformin (n=110)	n (%)	15 (13.8)	NR	0.53
		Insulin (n=107)		18 (16.8)	NR	
	Farrar 2016; 5 RCTs	Metformin	NR	NR	1.03 (0.66 to 1.62)	NR
		Insulin	NR	NR	1 (ref)	
Method of infant feeding						
Breastfeeding		Metformin, 7-year cohort (n=109)	n (%)	32 (55.1)	NR	NR

Outcome	Study	Intervention/Comparator	Outcome measure	Outcome value	RR (95% CI)	p-value
	Rowan 2018 (MiG)[99]	Insulin, 7-year cohort (n=51)		25 (49.0)	NR	
		Metformin, 9-year cohort (n=45)		25 (55.6)	NR	NR
		Insulin, 9-year cohort (n=54)		30 (56.6)	NR	
	Pellonpera 2016[100]	Metformin (=110)	Mean months (SD)	6.31 (4.00)	NR	NR
		Insulin (n=107)		6.59 (4.44)	NR	NR
Breastfeeding exclusively	Pellonpera 2016[100]	Metformin (=110)	Mean months (SD)	2.76 (2.37)	NR	NR
		Insulin (n=107)		2.58 (2.43)	NR	NR
Formula feeding	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109) Insulin, 7-year cohort (n=51)	n (%)	17 (29.3)	NR	NR
				13 (25.5)	NR	

		Metformin, 9-year cohort (n=45)		5 (11.1)	NR	NR
		Insulin, 9-year cohort (n=54)		10 (18.9)	NR	
Both breast and formula feeding	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109)	n (%)	5 (8.6)	NR	NR
		Insulin, 7-year cohort (n=51)		13 (25.5)	NR	
		Metformin, 9-year cohort (n=45)		14 (31.1)	NR	NR
		Insulin, 9-year cohort (n=54)		13 (24.5)	NR	
Long-term outcomes						
Post-pregnancy T2D	Pellonpera 2016[100]	Metformin (=110)	n (%)	4 (3.9)	NR	NR
		Insulin (n=107)		5 (5.0)	NR	NR

Abbreviations: CI, confidence interval; MiG, Metformin in Gestational Diabetes study; NR, not reported; RCT, randomised controlled trial; RR, relative risk; SD, standard deviation; T2D, type 2 diabetes

Neonatal outcomes

The same 3 studies (2 RCTs and the Farrar 2016 SLR including data on 3 to 9 RCTs depending on the outcome) that reported on maternal outcomes compared metformin and insulin and reported on at least 1 neonatal outcome (Table 28).[4, 99, 100]

In the MiG trial, there was no significant difference between metformin and insulin for birth weight (p=0.10 for 7-year cohort, p=0.69 for 9-year cohort).[99] Similarly, there was no significant difference between metformin and insulin for macrosomia (reported by Pellonpera 2016 [p=0.21] [100] and Farrar 2016 [RR 0.75, 95% CI 0.57 to 0.98]).[4] The incidence of LGA was reported to be significantly higher in those recieving metformin in the 7-year cohort of the MiG trial (20.7%), compared with insulin (5.9%; p=0.029).[99] However, no significant difference was detected in the longer 9-year cohort (11.1% vs 11.1%; p=1.00) in the same study,[99] or in an analysis based on 6 RCTs in Farrar 2016 (RR 0.81, 95% CI 0.62 to 1.05).[4] Farrar 2016 was the only study to report on neonatal hypoglycaemia for this treatment comparison, finding a lower risk in women treated with metformin compared with those treated with insulin (RR 0.71, 95% CI 0.51 to 0.98).[4] Farrar 2016 was also the only study to report on any form of birth injury, finding no significant between-arm difference in shoulder dystocia (RR 0.99, 95% CI 0.67 to 1.05).[4]

The other neonatal outcomes reported for metformin vs insulin were NICU admission and 5 minute Apgar score, both reported by Pellonpera 2016 and Farrar 2016. Results were consistent across both studies, with neither finding a significant difference for either NICU admission (Pellonpera 2016: 30.1% vs 36.4%, p=0.36;[100] Farrar 2016: RR 0.79, 95% CI 0.61 to 1.01[4]) or 5 minute Apgar score (Pellonpera 2016: 1.02% vs 0.98%, p=0.81;[100] Farrar 2016: RR 3.06, 95% CI 0.31 to 29.26[4]) for insulin compared to metformin.

Similar to maternal outcomes, metformin and insulin appear comparable for neonatal outcomes; however, having no studies reporting these outcomes in screen-detected women precludes drawing the same conclusion for this population.

Table 28. Neonatal outcomes reported by trials of metformin vs insulin for GDM

Outcome	Study	Intervention/Comparator	Outcome	Outcome value	Risk ratio (95% CI)	p-value
Glucose tolerance						
Birth weight	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109)	Mean g (SD)	3,481 (565)	NR	0.10
		Insulin, 7-year cohort (n=51)		3,324 (431)	NR	
		Metformin, 9-year cohort (n=45)		3,284 (563)	NR	0.69
		Insulin, 9-year cohort (n=54)		3,238 (542)	NR	
Macrosomia	Pellonpera 2016,[100] Huhtala 2018[103]	Metformin (n=110)	n (%)	5 (4.6)	NR	0.21
		Insulin (n=107)		10 (9.3)	NR	
	Farrar 2016; 9 RCTs[4]	Metformin	NR	NR	0.75 (0.57 to 0.98)	NR
		Insulin	NR	NR	1 (ref)	NR
LGA	Rowan 2018 (MiG)[99]	Metformin, 7-year cohort (n=109)	n (%)	**12 (20.7)**	NR	**0.029**
		Insulin, 7-year cohort (n=51)		**3 (5.9)**	NR	
		Metformin, 9-year cohort (n=45)		5 (11.1)	NR	1.00
		Insulin, 9-year cohort (n=54)		6 (11.1)	NR	
	Farrar 2016; 6 RCTs[4]	Metformin	NR	NR	0.81 (0.62 to 1.05)	NR
		Insulin	NR	NR	1 (ref)	NR
Neonatal hypoglycaemia	Farrar 2016; 7 RCTs[4]	Metformin	NR	NR	**0.71 (0.51 to 0.98)**	NR
		Insulin	NR	NR	**1 (ref)**	NR
Birth injury						
Shoulder dystocia	Farrar 2016; 3 RCTs[4]	Metformin	NR	NR	0.99 (0.67 to 1.05)	NR
		Insulin	NR	NR	NR	NR
Other outcomes						
NICU admission	Pellonpera	Metformin (n=110)	n (%)	33 (30.1)	NR	0.36

	2016,[100] Huhtala 2018[103]	Insulin (n=107)		39 (36.4)	NR	
	Farrar 2016[4]; 8	Metformin	NR	NR	0.79 (0.61 to 1.01)	
	RCTs	Insulin	NR	NR	NR	
Apgar score, 5 min	Pellonpera 2016,[100] Huhtala 2018[103]	Metformin (n=110)	Mean (SD)	8.80 (1.02)	NR	0.81
		Insulin (n=107)		8.85 (0.98)	NR	
	Farrar 2016[4]	Insulin	NR	NR	3.06 (0.31 to 29.26)	NR
		Metformin	NR	NR	1 (ref)	NR

Abbreviations: CI, confidence interval; MiG, Metformin in Gestational Diabetes study; LGA, large for gestational age; NICU, neonatal intensive care unit; NR, not reported; RCT, randomised controlled trial; RR, relative risk; SD, standard deviation

Glibenclamide/glyburide vs insulin

Two RCTs and 1 SLR (Farrar 2016) compared glibenclamide/glyburide and insulin.[4, 96, 97]

Maternal and pregnancy outcomes

Two RCTs (INDAO 2018 and GRACES 2017) and the Farrar 2016 SLR (including data on 1 to 4 RCTs depending on the outcome) compared glibenclamide/glyburide and insulin and reported on at least 1 maternal or pregnancy outcome (Table 29). The GRACES 2017 trial included women who had failed to achieve adequate glycaemic control on metformin monotherapy.[97] For all reported outcomes for this comparison, differences between arms were either not statistically quantified, or there was no significant difference.

No study reported on gestational hypertension. Farrar 2016 was the only study to report on pre-eclampsia for this comparison and found neither treatment to confer significantly lower risk than the other (RR 1.14, 95% CI 0.60 to 2.18 for glibenclamide vs insulin).[4] The frequency of pre-term birth was similar between glyburide and insulin arms in the INDAO trial (glyburide 6.8% vs insulin 4.1%)[96] and though substantially different in the GRACES trial (glibenclamide 0% vs insulin 30%),[97] this may be explained by the small patient numbers (n=13 in glibenclamide, n=10 in insulin). In addition, the Farrar 2016 SLR did not find a clear benefit for either treatment (RR 0.50, 95% CI 0.05 to 5.24 for glyburide vs insulin) on the basis of 1 included RCT.[4] The INDAO trial also reported on the rates of several different modes of delivery. Rates of spontaneous vaginal delivery (55.9% vs 56.8%), assisted vaginal delivery (17.2% vs 15.2%) and C-section (emergency: 17.2% vs 13.1%; elective: 9.8% vs 14.9%) were similar across arms for glyburide/glibenclamide vs insulin (p values not reported). However, the rates of different modes of birth were statistically different (p=0.08 [data not shown]).[96]

The GRACES trial reported on change in maternal weight between randomisation and 36 weeks of gestation, seeing a mean increase of 1.8 ± 3.5 kg with glibenclamide and 1.0 ± 1.5 kg with insulin (mean difference –0.77, 95% CI –3.55 to 2.01 kg).[97] The GRACES trial also reported on rates of different modes of delivery, however as with pre-term birth, the results are unreliable due to small patient numbers.[97]

Glybenclamide/glyburide and insulin appear comparable for maternal and pregnancy outcomes. Nevertheless, the studies did not report any outcomes in screen detected women and so it is uncertain whether the same conclusion could be drawn for this population.

Table 29. Maternal and pregnancy outcomes reported by trials of glibenclamide/glyburide vs insulin

Outcome	Study	Intervention (n in arm)	Outcome measure	Outcome value	RR (95% CI)	p value
Pregnancy outcomes						
Pre-eclampsia	Farrar 2016; 2 RCTs[4]	Glibenclamide	NR	NR	1.14 (0.60 to 2.18)	NR
		Insulin	NR	NR	1 (ref)	
Gestational age at birth	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	Median weeks (IQR)	38.3 (38.0 to 39.4)	Median difference –0.71 (–1.86 to 0.29)	NR
		Insulin (n=10)		38.1 (36.4 to 38.6)		
Preterm birth	INDAO (Senat 2018)[96]	Glyburide (n=367)	n (%)	25 (6.8)	NR	NR
		Insulin (n=442)		18 (4.1)	NR	NR
	Farrar 2016; 1 RCT[4]	Glyburide	NR	NR	0.50 (0.05 to 5.24)	NR
		Insulin	NR	NR	1 (ref)	NR
	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	0 (0.0)	NR	NR
		Insulin (n=10)		3 (30.0)	NR	
Gestational weight gain	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	Mean kg (SD)	1.8 (3.5)	Mean difference –0.77 (–3.55 to 2.01)	NR
		Insulin (n=10)		1.0 (1.5)		
Mode of birth						
Spontaneous vaginal	INDAO (Senat 2018)[96]	Glyburide (n=367)	n (%)	205 (55.9)	NR	NR
		Insulin (n=442)		251 (56.8)	NR	
	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	8 (61.5)	NR	NR
		Insulin (n=10)		3 (30.0)	NR	
Assisted/instrumental vaginal	INDAO (Senat 2018)[96]	Glyburide (n=367)	n (%)	63 (17.2)	NR	NR
		Insulin (n=442)		67 (15.2)	NR	
	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	1 (7.7)	NR	NR
		Insulin (n=10)		1 (10.0)	NR	
C-section	INDAO (Senat 2018)[96]	Glyburide (n=367)	n (%)	Emergency: 63 (17.2) Elective: 36 (9.8)	NR	NR
		Insulin (n=442)		Emergency: 58 (13.1) Elective: 66 (14.9)	NR	
	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	Emergency: 0 (0.0) Elective: 4 (30.8)	NR	NR
		Insulin (n=10)		Emergency: 4		

				(40.0) Elective:2 (20.0)	NR	
	Farrar 2016; 4 RCTs[4]	Glibenclamide	NR	NR	0.86 (0.66 to 1.12)	NR
		Insulin	NR	NR	1 (ref)	NR

Abbreviations: CI, confidence interval; GRACES, Glibenclamide and metfoRmin versus stAndard care in gEstational diabeteS; INDAO, Insulin Daonil; IQR, interquartile range; NR, not reported; RCT, randomised controlled trial; RR, relative risk; SD, standard deviation

Neonatal outcomes

The same 2 RCTs (INDAO 2018 and GRACES 2017) and the Farrar 2016 SLR (including data on 2 to 5 RCTs depending on the outcome) also compared glibenclamide/glyburide and insulin and reported on at least 1 neonatal outcome (Table 30).[4, 96, 97]

The Farrar 2016 SLR reported found no significant difference between glibenclamide vs insulin for either macrosomia (RR 2.66, 95% CI 0.91 to 7.77) or LGA (RR 2.44, 95% CI 0.97 to 6.15).[4] Similarly, they also found no significant difference in neonatal hypoglycaemia (RR 1.60, 95% CI 0.99 to 2.60).[4] The GRACES trial also reported on neonatal hypoglycaemia and found 3 events (27.3%) in the glibenclamide arm vs 1 event (11.1%) in the insulin arm, however no measures of statistical significance were reported and small patient numbers make it difficult to draw robust conclusions.[97]

The INDAO study of glyburide versus insulin found no significant difference between treatment arms for all birth injuries, including shoulder dystocia, bone fracture, nerve palsy and other injuries (p=0.66).[96] The INDAO trial also reported perinatal mortality. There were no deaths in the glyburide arm, and 2 deaths in the insulin arm, although no statistical analyses were conducted.[96]

There were no apparent differences for either severe respiratory distress syndrome or admission to NICU in the 2 and 3 RCTs reporting these outcomes, respectively (p>0.05 or not reported). The GRACES study also reported on the frequency of Apgar score <7 at 5 minutes of age, recording 0 events in both the glibenclamide and insulin arms.[97]

Similar to maternal outcomes, glybenclamide/glyburide and insulin appear comparable for neonatal outcomes; however, lack of studies in screen-detected women precludes drawing the same conclusion for this population.

Table 30. Neonatal outcomes reported by trials of glibenclamide/glyburide vs insulin for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	p-value
Glucose tolerance						

Macrosomia	Farrar 2016; 4 RCTs[4]	Glibenclamide	NR	NR	2.66 (0.91 to 7.77)	NR
		Insulin	NR	NR	1 (ref)	NR
LGA	Farrar 2016; 5 RCTs[4]	Glibenclamide	NR	NR	2.44 (0.97 to 6.15)	NR
		Insulin	NR	NR	NR	NR
Neonatal hypoglycaemia	Farrar 2016; 4 RCTs[4]	Glibenclamide	NR	NR	1.60 (0.99 to 2.60)	NR
		Insulin	NR	NR	1 (ref)	NR
	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	3 (27.3)	NR	NR
		Insulin (n=10)		1 (11.1)		
Outcome	**Study**	**Intervention/ Comparator**	**Outcome**	**Outcome value**	**Risk ratio (95% CI)**	**p-value**
Birth injury						
Shoulder dystocia	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	0 (0)	NR	NR
		Insulin (n=10)		0 (0		
	INDAO (Senat 2018)[96]	Glyburide (n=367)	n	1	NR	0.66 (for all birth injury, glyburide vs insulin)
		Insulin (n=442)		2	NR	
Bone fracture	INDAO (Senat 2018)[96]	Glyburide (n=367)	n	1	NR	
		Insulin (n=442)		6	NR	
Nerve palsy	INDAO (Senat 2018)[96]	Glyburide (n=367)	n	1	NR	
		Insulin (n=442)		0	NR	
Other	INDAO (Senat 2018)[96]	Glyburide (n=367)	n	3	NR	
		Insulin (n=442)		1	NR	
Other outcomes						
Perinatal mortality	INDAO (Senat 2018)[96]	Glyburide (n=367)	n	0	NR	NR
		Insulin (n=442)		2	NR	
Severe respiratory distress syndrome	INDAO (Senat 2018)[96]	Glyburide (n=367)	n (%)	8 (1.9)	NR	0.75
		Insulin (n=442)		11 (2.2)	NR	
	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	0 (0)	NR	NR
		Insulin (n=10)		0 (0		
NICU admission	INDAO (Senat 2018)[96]	Glyburide (n=367)	n (%)	Before 47h of life: 10 (2.3) Admission to neonatal ward: 27 (7.9)	NR	Before 47h of life: 0.87 Admission to neonatal ward: 0.86
		Insulin (n=442)		Before 47h of life: 11 (2.4) Admission to neonatal ward: 34 (8.2)	NR	
	Farrar 2016; 2 RCTs[4]	Glibenclamide	NR	NR	0.95 (0.49 to 1.84)	NR

		Insulin			NR	
	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	4 (30.8)	NR	NR
		Insulin (n=10)		1 (10.0)		
Apgar score <7 at 5 minutes of age	GRACES (Reynolds 2017)[97]	Glibenclamide (n=13)	n (%)	0 (0)	NR	NR
		Insulin (n=10)		0 (0		

Abbreviations: CI, confidence interval; GRACES, Glibenclamide and metfoRmin versus stAndard care in gEstational diabeteS; INDAO, Insulin Daonil; IQR, interquartile range; LGA, large for gestational age; NICU, neonatal intensive care unit; NR, not reported; RCT, randomised controlled trial; RR, relative risk; SD, standard deviation

Any oral antidiabetic agent vs insulin

Maternal and pregnancy outcomes

The Brown 2017I meta-analyses grouped all oral antidiabetic agents together for comparison with insulin for pre-eclampsia, induction of labour, C-section and postpregnancy type 2 diabetes (Table 31).[92] The results from these analyses largely supported those for the separate metformin or glyburide/glibenclamide comparisons with insulin. No significant differences were detected in the Brown 2017I meta-analysis of 10 RCTs comparing oral antidiabetic agents with insulin (RR 1.14, 95% CI 0.86 to 1.52).[92] Furthermore, the Brown 2017I analysis of 3 RCTs found no evidence of a reduced risk of induced labour between treatment with an oral antidiabetic agent versus insulin (RR 1.30, 95% CI 0.96 to 1.75).[92] This was consistent with the Farrar 2016 finding of no significant difference for metformin vs insulin,[4] but in contrast with Pellonpera 2016 who found that induction of labour was significantly more common in those women treated with insulin rather than metformin (Table 28).[81]

With respect to C-sections, Brown 2017I found no significant difference in the risk between insulin versus oral antidiabetic agent across 17 RCTs (RR 1.03, 95% CI 0.93 to 1.14), supporting findings of Pellonpera 2016 and Farrar 2016.[92] The Brown 2017I SLR also found no significant difference between comparisons of insulin vs oral antidiabetic agents for postpregnancy type 2 diabetes (2 RCTs, RR 1.39, 0.80 to 2.44),[92] adding weight to the result from Pellonpera 2016,[81] where rates appeared similar across metformin and insulin arms but the level of significance was not confirmed.

In summary, maternal and pregnancy outcomes were comparable between insulin and antidiabetic agents, but it is not known whether this conclusion can be extended to screendetected women, due to lack of reporting for this group.

Table 31. Maternal and pregnancy outcomes reported by trials of oral antidiabetic agent vs insulin

Outcome	Study	Intervention (n in arm)	Outcome measure	Outcome value	RR (95% CI)	pvalue

Pregnancy outcomes						
Preeclampsia	Brown 2017I (10 RCTs)[92]	Oral antidiabetic agent	Anticipated absolute effect (risk)[a] (95%CI)	77 per 1000	1 (ref)	NR
		Insulin		88 per 1000 (95% CI 66 to 117)	1.14 (0.86 to 1.52)	
Mode of birth						
Induction of labour	Brown 2017I (3 RCTs)[92]	Oral antidiabetic agent	Anticipated absolute effect (risk)[a] (95%CI)	408 per 1000	1 (ref)	NR
		Insulin		535 per 1000 (95% CI 424 to 669)	RR 1.30 (95% CI 0.96 to 1.75)	
C-section	Brown 2017I (17 RCTs)[92]	Oral antidiabetic agent	Anticipated absolute effect (risk)[a] (95%CI)	394 per 1000	1 (ref)	NR
		Insulin		405 per 1000 (95% CI 366 to 449)	1.03 (0.93 to 1.14)	
Long-term outcomes						
Postpregnancy T2D	Brown 2017I (2 RCTs)[92]	Oral antidiabetic agent	Anticipated absolute effect (risk)[a] (95%CI)	52 per 1000	1 (ref)	NR
		Insulin		73 per 1000 (95% CI 42 to 128)	1.39 (0.80 to 2.44)	

Abbreviations: CI, confidence interval; NR, not reported; RCT, randomised controlled trial; RR, relative risk; T2D, type 2 diabetes

Neonatal outcomes

Brown 2017I also reported on several neonatal outcomes for the comparison between any oral antidiabetic agent and insulin (Table 32). There was no significant difference in risk of macrosomia (13 RCTs, RR 1.01, 95% CI 0.76 to 1.35) or neonatal hypoglycaemia (24 RCTs, RR 1.14, 95% CI 0.85 to 1.52) between treatment with oral antidiabetics and insulin during pregnancy. There was also no evidence of a difference in risk of perinatal mortality between women treated with an oral antidiabetic agents and women treated with insulin, based on 10 RCTs (RR 0.85, 95% 0.29 to 2.49).[92] Since Brown 2017I was not specifically reporting on screen-detected women, it is unclear whether the interventions would also be comparable in that group.

Table 32. Neonatal outcomes reported by trials of oral antidiabetic agents vs insulin for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Glucose tolerance						
Macrosomia	Brown 2017I (13 RCTs)[92]	Oral antidiabetic agent	Anticipated absolute effect (risk)[a]	159 per 1000	1 (ref)	NR
		Insulin		161 per 1000 (95% CI 121 to 215)	1.01 (0.76 to 1.35)	
	Brown	Oral antidiabetic agent		111 per 1000	1 (ref)	NR

Neonatal hypoglycaemia	2017I (24 RCTs)[92]	Insulin	Anticipated absolute effect (risk)[a]	126 per 1000 (95% CI 94 to 169)	1.14 (0.85 to 1.52)	
Other neonatal outcomes						
Outcome	**Study**	**Intervention/ Comparator**	**Outcome**	**Outcome value**	**Risk ratio (95% CI)**	**pvalue**
Perinatal mortality	Brown 2017I (10 RCTs)[92]	Oral antidiabetic agent	Anticipated absolute effect (risk)[a]	8 per 1000	1 (ref)	NR
		Insulin	Anticipated absolute effect (risk)[a]	7 per 1000 (95% CI 2 to 20)	RR 0.85 (0.29–2.49)	

Abbreviations: CI, confidence interval; NR, not reported; RCT, randomised controlled trial; RR, relative risk

Glibenclamide/glyburide vs metformin

Maternal and pregnancy outcomes

The Brown 2017A SLR compared glibenclamide with metformin and reported on preeclampsia, induction of labour and C-section in analyses based on 2, 1 and 4 RCTs, respectively (Table 33).[93] There was no clear benefit of 1 treatment over the other for any of the outcomes (pre-eclampsia: RR 0.70, 95% CI 0.38 to 1.30; induction of labour: RR 0.81, 95% CI 0.61 to 1.07; C-section: RR 1.20, 95% CI 0.83 to 1.72, for metformin vs glibenclamide).[93]

Table 33. Maternal and pregnancy outcomes reported by trials of glibenclamide/glyburide vs metformin for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Pregnancy outcomes						
Preeclampsia	Brown 2017A (2 RCTs)[93]	Glibenclamide	Anticipated absolute effects	88 per 1000	1 (ref)	NR
		Metformin		62 per 1000 (95% CI 33 to 114)	0.70 (0.38 to 1.30)	
Mode of delivery						
Induction of labour	Brown 2017A (1 RCT)[93]	Glibenclamide	Anticipated absolute effects	613 per 1000	1 (ref)	NR
		Metformin		496 per 1000 (95% CI 374 to 655)	0.81 (0.61 to 1.07)	
C-section	Brown 2017A (4 RCTs)[93]	Glibenclamide	Anticipated absolute effects	392 per 1000	1 (ref)	NR
		Metformin		470 per 1000 (95% CI 325 to 674)	1.20 (0.83 to 1.72)	

Abbreviations: CI, confidence interval; NR, not reported; RCT, randomised controlled trial; RR, relative risk

Neonatal outcomes

Brown 2017A and Farrar 2016 compared glibenclamide with metformin and reported on several neonatal outcomes (Table 34). Neither treatment was evidently superior to the other for any outcome, with large confidence intervals for reported RRs.[4, 93]

Table 34. Neonatal outcomes reported by trials of glibenclamide/glyburide vs metformin for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Glucose tolerance						
Macrosomia	Farrar 2016; 1 RCT[4]	Glibenclamide	NR	NR	4.05 (0.46 to 35.42)	NR
		Metformin		NR	1 (ref)	
LGA	Farrar 2016; 1 RCT[4]	Glibenclamide	NR	NR	2.29 (1.09 to 4.81)	NR
		Metformin		NR	1 (ref)	
	Brown 2017A (2 RCTs)[93]	Glibenclamide	Anticipated absolute effects	193 per 1000	1 (ref)	NR
		Metformin		129 per 1000 (95% CI 46 to 354)	0.67 (0.24 to 1.83)	
Neonatal hypoglycaemia	Brown 2017A (4 RCTs)[93]	Glibenclamide	Anticipated absolute effects	48 per 1000	1 (ref)	NR
		Metformin		41 per 1000 (95% CI 20 to 84)	0.86 (0.42 to 1.77)	
	Farrar 2016; 2 RCTs[4]	Glibenclamide	NR	NR	1.19 (0.57 to 2.48)	NR
		Metformin	NR	NR	1 (ref)	
Birth injury						
Shoulder dystocia	Farrar 2016; 1 RCT[4]	Glibenclamide	NR	NR	3.04 (0.13 to 73.44)	NR
		Metformin		NR	1 (ref)	
Other outcomes						
NICU admission	Farrar 2016; 2 RCTs[4]	Glibenclamide	NR	NR	0.69 (0.29 to 1.66)	NR
		Metformin		NR	1 (ref)	
	Farrar 2016; 1 RCT[4]	Glibenclamide	NR	NR	Mean difference 0.06 (95% CI −0.53 to 0.65)	NR
		Metformin		NR	1 (ref)	

Abbreviations: CI, confidence interval; LGA, large for gestational age; NICU, neonatal intensive care unit NR, not reported; RCT, randomised controlled trial; RR, relative risk

Lifestyle intervention vs usual care

Maternal and pregnancy outcomes

Four RCTs (MFMU Network RCT,[94] CHOICE diet study,[95] Trout 2016[104] and Kokic 2018[98]) and 2 SLRs (Farrar 2016[4] and Brown 2017L[91]) that compared a lifestyle intervention with usual care reported on at least 1 maternal or pregnancy outcome (Table 35). While 3 of the RCTs reported on a dietary intervention, Kokic 2018 investigated an exercise intervention.[98]

In the only study that specified that women had screen-detected GDM, the MFMU Network RCT, there was no evidence of a difference in gestational hypertension or pre-eclampsia between nutritional counselling and usual care, whether the intervention was initiated at 24 to 26 or 27 to 29 weeks' gestation (p=0.91).[94] This finding was reinforced by the Brown 2017L MA, as they found no significant difference in the risk of pregnancy-induced hypertension or in pre-eclampsia between lifestyle intervention and usual care.[91] However, when examining pre-eclampsia alone, Farrar 2016 found a lower risk in women receiving a dietary modification compared with usual care (RR 0.58, 95% CI 0.36 to 0.93). There were no reported significant differences between lifestyle intervention or usual care for gestational age at birth (reported by 3 studies),[95, 98, 104] pre-term birth (reported by 1 study)[4] or gestational weight gain (reported by 1 study)[95] (Table 35).

Lifestyle interventions during pregnancy also did not appear to significantly reduce the risk of C-section based on the identified evidence of 3 primary RCTs (p>0.05 in 2 studies, not reported in 1 study),[94, 95, 104] the Farrar 2016 analyses of 8 RCTs (RR 0.86, 95% CI 0.77 to 0.95)[4] and the Brown 2017L analyses of 10 RCTs (RR 0.90, 95% CI 0.78 to 1.05).[91] A similar trend of no difference was seen for induction of labour in 1 RCT comparing a lower carbohydrate diet with a usual pregnancy diet (35.3% vs 34.4%, p=0.94),[104] along with Farrar 2016 (4 RCTs, RR 1.12, 95% CI 0.82 to 1.52)[4] and Brown 2017L (4 RCTs, RR 1.20, 95% CI 0.99 to 1.46) comparing dietary interventions with usual care.[91]

There was also no evidence of a difference in instrumental delivery (5.56% vs 0%, p=0.784), prolonged labour (5.56% vs 10%, p=0.633) or induction of labour (11.11% vs 35%, p=0.346) between a structured exercise intervention compared with usual care (nutrition therapy) in Kokic 2018, the 1 study that investigated exercise as an intervention.[98]

Brown 2017L found no clear reduction in risk with lifestyle intervention for post-pregnancy type 2 diabetes (RR 0.98, 95% CI 0.54 to 1.76).[91] One trial found the risk of postnatal depression was 83 per 1000 (95% CI 53 to 132) in women treated with a lifestyle intervention, compared with usual care (169 per 1000), producing a statistically significant RR of 0.49 (95% 0.31 to 0.78). No studies that evaluated the impact of lifestyle interventions on method of infant feeding were identified in this review.

In summary, the only study performed specifically in screen-detected women did not find any maternal or pregnancy outcomes to be significantly better in dietary intervention

compared with usual care. Based on the MFMU study, it does not appear that dietary intervention would be beneficial compared with the standard of care.

Table 35. Maternal and pregnancy outcomes reported by trials of lifestyle interventions for GDM

Outcome	Study	Intervention	Outcome measure	Outcome value(s)	RR (95% CI)	p-value
Pregnancy outcomes						
Gestational hypertension or pre-eclampsia	MFMU Network RCT (Palatnik 2015, Casey 2015) [94, 102]	Nutrition counselling and diet therapy, 24–26 weeks (n=69)	n (%)	7 (10.3)	NR	0.91
		Usual care, 24–26 weeks (n=43)		6 (14.0)	NR	
		Nutrition counselling and diet therapy, 27–29 weeks (n=288)		26 (9.0)	NR	
		Usual care, 27–29 weeks (n=282)		37 (13.1)	NR	
	Brown 2017L [91]	Lifestyle intervention	Anticipated absolute effects (risk)[a] (95%CI)	90 per 1000 (51 to 157)	0.70 (0.40 to 1.22)	NR
		Usual care		129 per 1000 (NR)	1.00 (ref)	
Pre-eclampsia	Farrar 2016; 5 RCTs [4]	Diet modification	NR	NR	**0.58 (0.36 to 0.93)**	NR
		Usual care			**1 (ref)**	
Gestational age at birth	CHOICE diet study (Hernandez 2016) [95]	CHOICE diet (n=6)	n (%)	40.5 (0.5)	NR	NR
		LC/CONV diet (n=6)		39.2 (0.4)	NR	
	Kokic 2018 [98]	Structured exercise programme (n=18)	Mean weeks (SD)	38.89 (0.90)	NR	0.063
		Usual care (n=20)		39.45 (0.60)	NR	NR
	Trout 2016 [104]	Lower-carbohydrate diet (n=37)	Mean weeks (SD)	37.78 (1.66)	NR	0.96
		Usual pregnancy diet (n=31)		37.76 (1.74)	NR	
Pre-term birth	Farrar 2016; 4 RCTs [4]	Diet modification	NR	NR	0.75 (0.46 to 1.21)	NR
		Usual care	NR	NR	1 (ref)	
Gestational weight gain	CHOICE diet study (Hernandez 2016) [95]	CHOICE diet (n=6)	Mean kg (SD)	2.3 (1.2)	NR	NR
		LC/CONV diet (n=6)		1.7 (1.6)		
Mode of birth						

Spontaneous vaginal	NR	NR	NR	NR	NR	NR
		Lifestyle intervention		252 per 1000 (220 to	1.20 (0.99 to 1.46)	NR
Induction of labour	Brown 2017L (4 RCTs)[91]		Anticipated absolute effects (risk)[a] (95%CI)	285)		
		Usual care		211 per 1000	1.00 (ref)	
	Kokic 2018[98]	Structured exercise programme (n=18)	n (%)	3 (11.11)	NR	0.346
		Usual care (n=20)		7 (35)	NR	
	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	35.3	NR	0.94
		Usual pregnancy diet (n=31)		34.4	NR	
		Diet modification	NR	NR	1.12 (0.82 to 1.52)	NR

Outcome	Study	Intervention	Outcome measure	Outcome value(s)	RR (95% CI)	p-value
	Farrar 2016; 4 RCTs[4]	Usual care			1 (ref)	
Prolonged labour	Kokic 2018[98]	Structured exercise programme (n=18)	n (%)	1 (5.56)	NR	0.633
		Usual care (n=20)		2 (10)	NR	
Instrumental delivery	Kokic 2018[98]	Structured exercise programme (n=18)	n (%)	1 (5.56)	NR	0.784
		Usual care (n=20)		0 (0)	NR	
	Farrar 2016; 1 RCT[4]	Diet modification	NR	NR	1.37 (0.20 to 9.27)	NR
		Usual care			1 (ref)	
C-section	CHOICE diet study (Hernandez 2016)[95]	CHOICE diet (n=6)	n	0	NR	NR
		LC/CONV diet (n=6)		2	NR	NR
	MFMU Network RCT (Palatnik 2015, Casey 2015)[94, 102]	Nutrition counselling and diet therapy, 24–26 weeks (n=69)	n (%)	23 (33.8)	NR	0.57
		Usual care, 24–26 weeks (n=43)		15 (34.9)	NR	

		Nutrition counselling and diet therapy, 27–29 weeks (n=288)		77 (26.7)	NR	
		Usual care, 27–29 weeks (n=282)		93 (33.0)	NR	
	Farrar 2016; 8 RCTs[4]	Diet modification	NR	NR	**0.86 (0.77 to 0.95)**	NR
		Usual care			**1 (ref)**	
	Brown 2017L (10 RCTs)[91]	Lifestyle intervention	Anticipated absolute effects (95% CI)	342 per 1000 (296 to 399)	0.90 (0.78 to 1.05)	NR
		Usual care		380 per 1000	1.00 (ref)	
Primary Csection	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	29.4	NR	0.34
		Usual pregnancy diet (n=31)		40.6	NR	
Trauma or injury						
Perineal trauma/tear	Brown 2017L[91]	Lifestyle intervention	Anticipated absolute effects (95% CI)	518 per 1000 (463 to 588)	1.04 (0.93 to 1.18)	NR
		Usual care		498 per 1000	1.00 (ref)	
Long-term outcomes						
Post-pregnancy T2M	Brown 2017L[91]	Lifestyle intervention	Anticipated absolute effects (95% CI)	81 per 1000 (45 to 146)	0.98 (0.54 to 1.76)	NR
		Usual care		83 per 1000	1.00 (ref)	
Postnatal depression	Brown 2017L[91]	Lifestyle intervention	Anticipated absolute effects (95% CI)	**83 per 1000 (53 to 132)**	**0.49 (0.31 to 0.78)**	NR
		Usual care		**169 per 1000**	**1.00 (ref)**	

Abbreviations: CHOICE, higher-complex carbohydrate/lower fat; CI, confidence interval; LC/CONV, low-carbohydrate/higher-fat; MFMU, Maternal Fetal Medicines Unit; NR, not reported; RCT, randomised controlled trial; RR, relative risk; SD, standard deviation; T2D, type 2 diabetes

Neonatal outcomes

The same 4 RCTs (MFMU Network RCT,[94] CHOICE diet study,[95] Trout 2016[104] and Kokic 2018[98]) and 2 SLRs (Farrar 2016[4] and Brown 2017L[91]) that compared a lifestyle intervention with usual care reported for maternal/pregnancy outcomes also reported on neonatal outcomes (Table 36).

While no difference in LGA between nutritional counselling and usual care was reported in the MFMU Network trial (p=0.36),[94] the Brown 2017L comparison of lifestyle interventions with usual care found evidence of a significantly reduced risk of LGA in women allocated to lifestyle interventions based on 6 RCTs (RR 0.60, 95% CI 0.50 to 0.71). This was supported by findings in Farrar 2016 (6 RCTs, RR 0.55, 95% CI 0.44 to 0.69).[4] A lower carbohydrate diet was found to have no impact on the risk of macrosomia compared with usual diet in the Trout 2016 study (p=0.93),[104] but diet modification was shown to reduce this compared with routine antenatal care in Farrar 2016 (9 RCTs, RR 0.47, 95% CI 0.45 to 0.60).[4]

The Kokic 2018 trial reported that newborns born to women who had underwent a structured exercise programme during pregnancy had significantly lower neonatal BMI than those in the usual care (nutrition therapy) arm (p=0.035), although this difference was small (13.96 vs. 13.21 kg/m^2, respectively).[98]

No significant difference was found for neonatal hypoglycaemia between lifestyle interventions and usual care (6 RCTs, RR 0.99, 95% CI 0.65 to 1.52) in the Brown 2017L analysis. On the contrary, Trout 2016 reported that incidence of neonatal hypoglycaemia was lower in babies born to women in the lower-carbohydrate diet arm (9.7%) compared with a usual pregnancy diet (26.9%), but this did not reach statistical significance (p=0.09).[104] Due to a small sample size, no cases of neonatal hypoglycaemia occurred in the Kokic 2018 study of structured exercise compared with usual care.[98]

There was no difference in shoulder dystocia in the Trout 2016 study of low-carbohydrate diet versus usual diet (2.9% vs 0%, respectively; p=0.25), while no bone fracture or nerve palsy events occurred in either arm.[104] In contrast with Trout 2016, Farrar 2016 reported a reduction in shoulder dystocia in women in diet modification arms compared to those receiving usual care (4 RCTs, RR 0.39, 95% CI 0.23 to 0.69).[4]

Of the remaining outcomes, there was no clear evidence of reduction in perinatal mortality for women in lifestyle intervention arms vs usual care (RR 0.09, 95% CI 0.01 to 1.70). There

was also no difference between lifestyle intervention and usual care for NICU admission in the MFMU Network RCT (p=0.55),[94] Trout 2016 (p=0.38)[104] or Farrar 2016

(RR 0.91, 95% CI 0.62 to 1.34),[4] or for Apgar score in the trial comparing exercise with usual care (1 minute score: p=0.828; 5 minute score: p=1.000).[98]

Given the lack of differences between dietary interventions and usual care, and other studies not specifically reporting on screen-detected women, it cannot be concluded that lifestyle interventions would be beneficial compared with usual care, in this population.

Table 36. Neonatal outcomes in women with GDM undergoing lifestyle interventions

Outcome	Study	Study arm	Outcome	Outcome value	Risk ratio (95% CI)	p-value
Glucose tolerance						
Birth weight	CHOICE diet study (Hernandez 2016)[95]	CHOICE diet (n=6)	Mean g (SD)	3,273.0 (104.0)	NR	NR
		LC/CONV diet (n=6)		3,421.0 (186.3)	NR	NR
	Trout 2016[104]	Lower-carbohydrate diet (n=37)	Mean g (SD)	3,409.53 (527.91)	NR	0.81
		Usual pregnancy diet (n=31)		3,377.28 (589.91)	NR	
Neonatal body mass	Kokic 2018[98]	Structured exercise programme (n=18)	Mean g (SD)	3514.45 (413.57)	NR	0.393
		Usual care (n=20)		3377.00 (494.27	NR	
Neonatal BMI	Kokic 2018[98]	Structured exercise programme (n=18)	Mean kg/m^2 (SD)	**13.96 (0.97)**	NR	**0.035**
		Usual care (n=20)		**13.21 (1.01)**	NR	
Adiposity	NR	NR	NR	NR	NR	NR
Macrosomia	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	11.8	NR	0.93
		Usual pregnancy diet (n=31)		12.5	NR	
	Farrar 2016; 9 RCTs[4]	Diet modification	NR	NR	**0.47 (0.45 to 0.60)**	NR
		Routine antenatal care	NR	NR	**1 (ref)**	NR
LGA	MFMU Network RCT (Palatnik 2015, Casey 2015)[94, 102]	Nutrition counselling and diet therapy, 24–26 weeks (N=69)	n (%)	8 (11.6)	NR	0.36
		Usual care, 24–26 weeks (N=43)		6 (14.0)	NR	
		Nutrition counselling and diet therapy, 27–29 weeks (N=288)		20 (6.9)	NR	
		Usual care, 27–29 weeks (N=282)		40 (14.2)	NR	
	Farrar 2016; 6 RCTs[4]	Diet modification	NR	NR	**0.55 (0.44 to 0.69)**	NR
		Usual care	NR	NR	**1 (ref)**	NR
	Brown 2017L; 6 RCTs[91]	Lifestyle intervention		113 per 1000 (95% CI 95 to 134)	**0.60 (0.50 to 0.71)**	NR

			Anticipated absolute effects (95% CI)	189 per 1000	**1 (ref)**	NR
		Usual care				
Neonatal hypoglycaemia	Brown 2017L; 6 RCTs[91]	Lifestyle intervention	Anticipated absolute effects (95% CI)	74 per 1000 (95% CI 49 to 114)	0.99 (0.65 to 1.52)	NR
		Usual care		75 per 1000	1 (ref)	NR
	Kokic 2018[98]	Structured exercise programme (n=18)	n (%)	0 (0)	NR	1.00
		Usual care (n=20)		0 (0)	NR	
	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	**9.7**	NR	0.09
		Usual pregnancy diet (n=31)		**26.9**	NR	
Birth injury						
Shoulder dystocia	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	2.9	NR	0.25
		Usual pregnancy diet (n=31)		0	NR	
		Diet modification	NR	NR	**0.39 (0.23 to 0.69)**	NR

	Farrar 2016; 4 RCTs[4]	Usual care	NR	NR	**1 (ref)**	NR
Bone fracture	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	0	NR	NR
		Usual pregnancy diet (n=31)		0	NR	NR
Nerve palsy	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	0	NR	NR
		Usual pregnancy diet (n=31)		0	NR	NR
Other neonatal outcomes						
Perinatal mortality	Brown 2017L[91]	Lifestyle intervention	Anticipated absolute effects (95% CI)	0 per 1000 (95% CI 0 to 9)	0.09 (0.01 to 1.70)	NR
		Usual care		5 per 1000 (NR)	1.00 (ref)	NR
NICU admission	MFMU Network RCT (Palatnik 2015, Casey	Nutrition counselling and diet therapy, 24–26 weeks (n=69)	NR	10 (14.5)	NR	0.55
		Usual care, 24–26 weeks (n=43)	NR	7 (16.3)	NR	

	2015)[94, 102]	Nutrition counselling and diet therapy, 27–29 weeks (n=288)	NR	25 (8.7)	NR	
		Usual care, 27–29 weeks (n=282)	NR	38 (13.5)	NR	
	Trout 2016[104]	Lower-carbohydrate diet (n=37)	%	20.6	NR	0.38
		Usual pregnancy diet (n=31)		12.5	NR	
	Farrar 2016; 4 RCTs[4]	Diet modification	NR	NR	0.91 (0.62 to 1.34)	NR
		Usual care	NR	NR	1 (ref)	NR
Apgar score						
1 min score	Kokic 2018[98]	Structured exercise programme (n=18)	Mean (SD)	9.89 (0.47)	NR	0.828
		Usual care (n=20)		9.80 (0.70)	NR	
5 min score	Kokic 2018[98]	Structured exercise programme (n=18)	Mean (SD)	10.0 (0)	NR	1.000
		Usual care (n=20)		10 (0)	NR	

Abbreviations: BMI, body mass index; CHOICE, higher-complex carbohydrate/lower fat; CI, confidence interval; LC/CONV, low-carbohydrate/higher-fat; LGA, large for gestational age; MFMU, Maternal Fetal Medicines Unit; NICU, neonatal intensive care unit; NR, not reported; RCT, randomised controlled trial; RR, relative risk; SD, standard deviation

Insulin vs diet/standard care

Maternal and pregnancy outcomes

The Brown 2017I SLR reported on 4 maternal or pregnancy outcomes for the comparison of insulin with diet or standard care, with analyses including 1 or 2 RCTs in each case.[92] No significant difference was observed between arms for any of gestational age at birth (p=0.073), pre-term birth (RR 1.09, 95% CI 0.64 to 1.85; p=0.76), C-section (RR 0.85, 95% CI 0.50 to 1.42; p=0.53) or development of maternal type 2 diabetes after pregnancy (RR 0.98, 95% CI 0.79 to 1.21; p=0.83) (Table 37).[92]

Table 37. Maternal and pregnancy outcomes reported by trial of insulin vs diet or standard care for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Pregnancy outcomes						
Gestational age at birth	Brown 2017I (2 RCTs)[92]	Diet/standard care (n=45)	Mean difference, weeks (95% CI)	1 (ref)	NR	0.073
		Insulin (n=61)		–0.66 (–1.37 to 0.06)		
Pre-term birth	Brown 2017I (1 RCT)[92]	Diet/standard care (n=306)	n	24	1 (ref)	0.76
		Insulin (n=305)		26	1.09 (0.64 to 1.85)	
Mode of birth						
C-section	Brown 2017I (2 RCTs)[92]	Diet/standard care (n=61)	n	20	1 (ref)	0.53
		Insulin (n=72)		19	0.85 (0.50 to 1.42)	
Long-term maternal outcomes						
Maternal T2D	Brown 2017I (2 RCTs)[92]	Diet/standard care (n=319)	n	110	1 (ref)	0.83
		Insulin (n=334)		110	0.98 (0.79 to 1.21)	

Abbreviations: CI, confidence interval; NR, not reported; RCT, randomised controlled trial; T2D, type 2 diabetes

Neonatal outcomes

The Brown 2017I SLR reported on 7 neonatal outcomes for the same comparison of insulin and diet/standard care.[92] The only significant result was for macrosomia (RR 0.30, 95% CI 0.18 to 0.50; p<0.001 for insulin vs diet or standard care). No significant differences were observed for LGA, neonatal hypoglycaemia, perinatal or neonatal mortality (p>0.05) (Table 38). No events occurred in either arm for should dystocia or nerve palsy, therefore the effect size could not be estimated.[92]

Table 38. Neonatal outcomes reported by trial of insulin vs diet or standard care for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Glucose tolerance						
LGA	Brown 2017I (1 RCT)[92]	Diet/standard care (n=105)	n	14	1 (ref)	0.67
		Insulin (n=97)		11	0.85 (0.41 to 1.78)	
Macrosomia		Diet/standard care (n=351)	n	53	**1 (ref)**	**<0.001**
Outcome	**Study**	**Intervention/ Comparator**	**Outcome**	**Outcome value**	**Risk ratio (95% CI)**	**pvalue**
	Brown 2017I (3 RCTs)[92]	Insulin (n=366)		17	**0.30 (0.18 to 0.50)**	
Neonatal hypoglycaemia	Brown 2017I (3 RCTs)[92]	Diet/standard care (n=75)	n	18	1 (ref)	0.78
		Insulin (n=101)		22	0.88 (0.34 to 2.24)	
Birth trauma						
Shoulder dystocia	Brown 2017I (2 RCTs)[92]	Diet/standard care (n=61)	n	0	Not estimable	NR
		Insulin (n=72)		0		
Nerve palsy	Brown 2017I (1 RCT)[92]	Diet/standard care (n=11)	n	0	Not estimable	NR
		Insulin (n=27)		0		
Other neonatal outcomes						
Perinatal and later infant mortality	Brown 2017I (4 RCTs)[92]	Diet/standard care (n=579)	n	25	1 (ref)	0.31
		Insulin (n=558)		18	0.74 (0.41 to 1.33)	
Neonatal mortality	Brown 2017I (1 RCT)[92]	Diet/standard care (n=306)	n	7	1 (ref)	0.57
		Insulin (n=305)		5	0.72 (0.23 to 2.23)	

Abbreviations: CI, confidence interval; LGA, large for gestational age; NR, not reported; RCT, randomised controlled trial

Insulin vs exercise

The Brown 2017I SLR also included 1 RCT that reported on the comparison between insulin and exercise. This was conducted in the US and had 34 participants. The type of insulin used was not prespecified.[92]

There was no reported significant difference between any maternal outcomes (gestational age at birth [p=0.21] and C-section [p=0.63]) or neonatal outcomes (macrosomia [p=0.38], neonatal hypoglycaemia [p=0.56]) between women on the exercise regime compared to

those treated with insulin. No events occurred in either arm for respiratory distress syndrome so it was not possible to estimate the effect (Table 39 and Table 40).[92]

Table 39. Maternal and pregnancy outcomes reported by trial of insulin vs exercise for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Pregnancy outcomes						
Gestational age at birth	Brown 2017I (1 RCT)[92]	Exercise (n=17)	Mean difference, weeks (95% CI)	1 (ref)	NR	0.21
		Insulin (n=17)		–0.80 (–2.05 to 0.45)		
Mode of birth						
C-section	Brown 2017I (1 RCT)[92]	Exercise (n=17)	n	2	1 (ref)	0.63
		Insulin (n=17)		3	1.5 (0.29 to 7.87)	

Abbreviations: CI, confidence interval; NR, not reported; RCT, randomised controlled trial

Table 40. Neonatal outcomes reported by trial of insulin vs exercise for GDM

Outcome	Study	Intervention/ Comparator	Outcome	Outcome value	Risk ratio (95% CI)	pvalue
Glucose tolerance						
Macrosomia	Brown 2017I (1 RCT)[92]	Exercise (n=17)	n	2	1 (ref)	0.38
		Insulin (n=17)		4	2.0 (0.42 to 9.50)	
Neonatal hypoglycaemia	Brown 2017I (1 RCT)[92]	Exercise (n=17)	n	2	1 (ref)	0.56
		Insulin (n=17)		1	0.5 (0.05 to 5.01)	
Other neonatal outcomes						
Respiratory distress syndrome	Brown 2017I (1 RCT)[92]	Exercise (n=17)	n	0	Not estimable	NR
		Insulin (n=17)		0		

Abbreviations: CI, confidence interval; NR, not reported; RCT, randomised controlled trial

Conclusions

A summary of the available evidence for different treatment comparisons and maternal/pregnancy or neonatal outcomes is presented in Table 41. Evidence was identified for 7 different treatment comparisons:

[1] glibenclamide vs placebo
[2] metformin vs insulin;
[3] glibenclamide/glyburide vs insulin;

[4] any oral antidiabetic vs insulin;
[5] glibenclamide/ glyburide vs metformin; [6] lifestyle intervention vs usual care; [7] insulin vs lifestyle intervention.

Only 1 study specified that the included population was screen-detected GDM, with the rest all appearing to include clinically diagnosed GDM or not specifying any further details.

In the study comparing glibenclamide with placebo there was no evidence that treatment with glibenclamide significantly improved maternal or neonatal outcomes, including preeclampsia, induction of labour, C-section, LGA and neonatal hypoglycaemia.

Based on this review's findings, there is little evidence to suggest that oral pharmacological interventions such as metformin or glyburide given during pregnancy in women with clinically-diagnosed GDM are superior to insulin in reducing the risk of adverse pregnancy and postnatal outcomes. There was also no significant difference between glibenclamide/glyburide and metformin for any of the reported outcomes (pre-eclampsia, induction of labour, C-section, macrosomia, LGA, neonatal hypoglycaemia, shoulder dystocia or NICU admission), indicating that neither of these treatments is superior to the other for preventing these outcomes. Further research to explore this and the comparison of other antidiabetics with placebo or usual care could be useful.

For lifestyle interventions comprising a form of dietary modification, there was more evidence to suggest some differences between this and usual care (specific details of usual care varied by study and were not always reported but often included nutritional counselling). However, results were not consistent across multiple studies reporting on the same outcome. The 1 study reporting on pre-eclampsia and the 1 study reporting on postnatal depression found the risks were lower in women receiving diet modification, however, without replication in other studies, it is difficult to evaluate the robustness of these results. At least 1 study reporting on C-section, macrosomia, LGA, neonatal hypoglycaemia and shoulder dystocia found lower risks for women receiving dietary modification compared with usual care, but all of these outcomes were also reported by at least 1 study that found no significant differences. The 1 trial that reported on an exercise intervention compared with usual care only reported 1 significant result: neonatal BMI was significantly lower for exercise vs usual care. The only significant result for insulin vs lifestyle intervention was for macrosomia in insulin compared to diet/standard care.

Lacking data and/or poor reporting of statistical differences means that no conclusions can be drawn for any treatment comparisons for spontaneous vaginal delivery, method of infant feeding, bone fracture or nerve palsy.

Overall, the substantial volume of evidence identified, including that from 4 high quality SLRs, does not suggest that there is any treatment that is clearly superior to the other for any of the treatment comparisons identified for women with clinically diagnosed GDM. Importantly, evidence is lacking in 2 key areas. The first of these is a lack of comparison between interventions and placebo or usual care. The majority of evidence compared two interventions, with only 1 SLR comparing glibenclamide with placebo, and a limited number of studies comparing lifestyle intervention with usual care. However, the benefit of interventions examined in this review against no treatment has been demonstrated previously, most notably by the ACHOIS study. The second evidence gap is the lack of studies clearly reporting on a population of women with screen-detected GDM. This was only reported by 1 study identified in the rapid review. Therefore, whilst it may be possible to make assumptions based on treatment effects in clinically-diagnosed populations, the effect of treatment for the screen-detected population remains highly uncertain.

Table 41. Summary of the number of studies and direction of evidence for each outcome

Outcome	Metformin vs insulin	Glibenclamide/ glyburide vs insulin	Any oral antidiabetic vs insulin	Glibenclamide/ glyburide vs metformin	Glibenclamide vs placebo	Lifestyle intervention vs usual care	Insulin vs lifestyle intervention
	N studies Comparison	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**
Any maternal or pregnancy outcome	**1 SLR 2 RCTs**	**1 SLR 2 RCTs**	**1 SLR**	**1 SLR**	**1 SLR**	**2 SLRs 4 RCTs**	**1 SLR**
Gestational hypertension	2 No sig. dif.	0 NA	0 NA	0 NA	0 NA	0 NA	0 NA
Pre-eclampsia	3 No sig. dif.	1 No sig. dif.	1 No sig. dif.	1 No sig. dif.	1 No sig. dif.	1 Lower for diet modification	0 NA
Gestational age at birth	2 Longer for insulin in 1 study; no sig. dif. in 1 study	1 NR	0 NA	0 NA	0 NA	3 No sig. dif.	1 No sig. dif.
Pre-term birth	2 No sig. dif.	3 No sig. dif.	0 NA	0 NA	0 NA	1 No sig. dif.	1 No sig. dif.
Gestational weight gain	0 NA	1 No sig. dif.	0 NA	0 NA	0 NA	1 NR	0 NA
Spontaneous vaginal delivery	0 NA	2 NR	0 NA	0 NA	0 NA	0 NA	0 NA
Assisted/instrumental vaginal delivery	2 Higher for metformin in 1 study; no sig. dif. In 1 study	2 NR	0 NA	0 NA	0 NA	2 No sig. dif.	0 NA
Induction of labour	2 Higher for insulin in 1 study; no significant difference in 1 study	0 NA	1 No sig. dif.	1 No sig. dif.	1 No sig. dif.	4 No sig. dif.	0 NA
C-section	2 No sig. dif.	3 No sig. dif.	1 No sig. dif.	1 No sig. dif.	1 No sig. dif.	5 Lower for diet modification in 1 study; no sig. dif. in 2 studies; NR in 2 studies	1 No sig. dif.
Method of infant feeding	2 NR	0 NA	0 NA	0 NA	0 NA	0 NA	0 NA
Post-pregnancy type 2 diabetes	1 NR	0 NA	1 No sig. dif.	0 NA	0 NA	1 No sig. dif.	1 No sig. dif.

Post-natal depression	0 NA	0 NA	0 NA	0 NA	0 NA	1 Lower for lifestyle intervention	0 NA
Any neonatal outcome	**1 SLR 2 RCTs**	**1 SLR 2 RCTs**	**1 SLR**	**2 SLRs**	**1 SLR**	**2 SLRs 4 RCTs**	**1 SLR**
Birth weight	1 No sig. dif.	0 NA	0 NA	0 NA	0 NA	2 No sig. dif.	0 NA
Macrosomia	2 No sig. dif.	1 No sig. dif.	1 No sig. dif.	1 No sig. dif.	0 NA	2 Lower for diet modification in 1 study; no sig. dif. in 1 study	1 Lower for insulin vs diet; no sig. dif. for insulin vs exercise

Outcome	Metformin vs insulin	Glibenclamide/ glyburide vs insulin	Any oral antidiabetic vs insulin	Glibenclamide/ glyburide vs metformin	Glibenclamide vs placebo	Lifestyle intervention vs usual care	Insulin vs lifestyle intervention
	N studies Comparison	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**	**N studies Comparison**
LGA	2 Higher for metformin in 1 study; no sig. dif. in 1 study	1 No sig. dif.	0 NA	2 No sig. dif.	1 No sig. dif.	3 Lower for lifestyle intervention in 2 studies; no sig. dif. in 1 study	1 No sig. dif.
Neonatal hypoglycaemia	1 Lower for metformin	2 No sig. dif.	1 No sig. dif.	2 No sig. dif.	1 No sig. dif.	3 Lower for diet modification in 1 study; no sig. dif. in 2 studies (including 1 on exercise)	1 No sig. dif.
Shoulder dystocia	1 No sig. dif.	2 NR	0 NA	2 No sig. dif.	0 NA	2 Lower for diet modification in 1 study; no sig. dif. in 1 study	1 NA
Bone fracture	0 NA	1 NR	0 NA	0 NA	0 NA	1 NR	0 NA
Nerve palsy	0 NA	1 NR	0 NA	0 NA	0 NA	1 NR	1 NA
Perinatal mortality	0 NA	1 NR	1 No sig. dif.	0 NA	0 NA	1 No sig. dif.	1 No sig. dif.
Severe respiratory distress syndrome	0 NA	2 No sig. dif.	0 NA	0 NA	0 NA	0 NA	1 No sig. dif.
NICU admission	2 No sig. dif.	3 No sig. dif.	0 NA	1 No sig. dif.	0 NA	3 No sig. dif.	0 NA

Apgar score	1 No sig. dif.	1 NR	0 NA	0 NA	0 NA	1 No sig. dif.	0 NA

Summary of Findings Relevant to Criterion 9: Criterion not met‡

Quantity: Including the data in 4 included SLRs, a high volume of evidence was available to assess Criterion 9 for clinically diagnosed GDM, consisting of a total of 34 RCTs reported across 4 SLRs and 7 primary publications (Tier 1 evidence). However, only 1 RCT with a screendetected population of GDM was included. Evidence was identified for 7 different direct treatment comparisons (Table 41). The lifestyle intervention was predominantly dietary modification, with only 1 study reporting on an exercise programme. Although few primary studies reported on each treatment comparison, at least 1 SLR reported on each (with between 8 and 51 RCTs included in MAs). No evidence was identified for comparisons between antidiabetic agents (e.g. metformin or glibenclamide/glyburide) vs lifestyle interventions and there was limited evidence for the comparison of interventions with placebo or usual care – 1 SLR compared glibenclamide with placebo and 2 SLRs and 4 RCTs compared lifestyle interventions with usual care.

Quality: All 4 SLRs were judged to be at a low risk of bias. Two primary RCTs were at a high risk of bias due to missing outcome data and measurement of the outcomes. [95, 99] There were some concerns about the risk of bias in 3 RCTs, particularly for the effect of assignment to the interventions as a result of limited information surrounding allocation concealment, and selection of the reported result due to unavailability of protocols or statistical analysis plans. The remaining 3 trials were at a low risk of bias for all study domains.[96-98]

Applicability: The main concern regarding applicability arises from the lack of studies in screendetected women. All but 1 study was in women clinically diagnosed with GDM or populations whose origin (screening or clinical diagnosis) were not reported. Otherwise, there were concerns about applicability due to the inclusion of non-EEA or OECD countries in all 4 of the included

SLRs. In most cases, this is not judged to have too high an implication for applicability as ≥50% of the studies were located high income countries. The exception to this is the Brown 2017A SLR in which ≥50% of the included studies were located in non-EEA/OECD countries.

Consistency: All of the included SLRs provided satisfactory discussions of the heterogeneity of their included studies. There are also low concerns regarding the approach to data analyses conducted in different SLRs. However, where multiple studies reported on the same outcome for the same treatment comparison, there appeared to be lack of consistency in the results in that the same treatment effects were not seen across multiple studies. For several of the treatment

‡ **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.

Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

comparisons, conclusions about consistency of RCTs could not be drawn because only a single RCT was included.

Conclusions: Overall, the evidence did not support increased effectiveness of any specific intervention compared with another intervention, or compared with no treatment/placebo for improving outcomes in pregnant women with GDM, screen-detected or otherwise. Evidence was lacking in 2 key areas: 1) comparison of interventions with placebo or usual care; 2) studies including populations with screen-detected GDM. It is uncertain whether the conclusions based on clinically-detected GDM could be applied to a screen-detected population. Based on the lack of evidence, Criterion 9 is judged to be not met.

Review summary

Conclusions and implications for policy

Based on the overall synthesis of evidence identified in this rapid review against the UK NSC criteria, screening for GDM is still not recommended. GDM and hyperglycaemia are important health problems and there appear to be moderately safe treatments available. However, it is unclear whether benefits would outweigh the harms if universal screening for GDM were to be introduced. This is because of uncertainties around the thresholds at which women should be considered at risk, the lack of a safe and practical test or lack of data supporting the use of OGTT as a screening test, and lack of data supporting benefits from currently available interventions in screen-detected women.

Three questions were considered in this review: (1) what are the risks of adverse outcomes associated with incremental increases in blood glucose level in the newborn; (2) what are the most effective screening tests or strategies to identify women at risk of hyperglycaemia in pregnancy or GDM; and (3) what is the most effective intervention for lowering glucose levels in screen-detected pregnant women with GDM and preventing adverse perinatal outcomes? The aim was for questions to consider populations of women without risk factors who may develop GDM but would not be identified based on the current NICE risk factorbased screening approach.

A large number of studies examining the effect of increased blood glucose on pregnancy and newborn outcomes was identified, however, most studies considered thresholds where a GDM diagnosis would be made under the current NICE guidelines (though not under other guidelines). Only 2 studies included a group of women without risk factors where glucose was elevated yet still under the NICE threshold for GDM. In those studies, the 2 outcomes where risk was consistently higher than with normal glucose tolerance were LGA and macrosomia. This was further confirmed by the other studies included in the review that did not specifically include only women below the NICE threshold, but who nevertheless included women not considered as having had GDM at the time of their pregnancy. The implications of this are that women with elevated glucose appear to be at risk of at least some adverse outcomes. Although the outcomes were not presented in a way that would identify a threshold at which the risk becomes significant, thresholds used in those 2 studies were 5.1 mmol/L in FPG and 5.7 mmol/L with a 75 g OGTT, indicating that at these threshold there is an increased risk to the pregnancy. It may be that for the 2 other outcomes where an association was consistently found (C-section and induction of labour) the risk is also increased in low risk women with elevated glucose; however, as these outcomes were not investigated by any studies that included women without risk factors

where glucose was elevated but below the current NICE threshold for GDM, no conclusions could be drawn.

Criterion 1 was not met as whilst it is clear that hyperglycaemia increases the risk of at least certain pregnancy and neonatal outcomes, no clear glucose threshold at which risk becomes substantially increased could be identified. Notably, the risk of adverse outcomes appears to be a continuum, and it may be that no ideal threshold could be determined. Instead, a threshold encompassing the best balance between reduction in risk from GDM and avoidance of overtreatment could be sought.

No test has been found that could be used for screening a low risk population other than the OGTT. However, in the current NICE guidance, the OGTT is both the diagnostic test and the reference standard. As such, its own reliability can only be assessed by comparison to a clinical diagnosis. In addition, the OGTT includes glucose loading, which could be harmful for those with impaired glucose tolerance, i.e. the exact group of women it intends to identify. Moreover, there are possible side effects including nausea and vomiting, and practical implications, as the test needs to be taken over 2 hours, which may discourage some women from attending. However, based on the balance between sensitivity (when trying to limit the number of positive women who would be indicated for OGTT) and specificity (if trying to comprehensively identify all women with GDM) of alternative tests and strategies (including FPG, GCT, risk factors or biomarkers) OGTT alone remains the best currently available screening test. Screening with any other test before OGTT (in order to avoid OGTT/glucose loading) either misses GDM cases or still requires almost all women to undergo the OGTT. Without further data on the safety and acceptability of the OGTT, or availability of a better test, Criterion 4 is not met.

Criterion 9 is judged to be not met because despite a large evidence base, only 1 study included a confirmed screen-detected GDM population and few studies compared interventions with placebo or usual care. In clinically-diagnosed GDM, none of the interventions could be shown to be consistently better than the other. It is therefore likely that they are similarly effective. While their benefit over no treatment is not certain, the benefit of interventions examined in this review against no treatment has been demonstrated previously, most notably by the ACHOIS study. Criterion 9 was specifically determined to be not met because studies have not demonstrated that the interventions are of benefit when applied to women who are screen-detected rather than those who are clinically diagnosed. In most studies, the basis upon which women were investigated for GDM was unknown; women could have been referred for diagnosis based on presence of risk factors or clinical symptoms, or the finding of GDM may have been incidental. Therefore, it is uncertain whether the results of interventions in these groups are similar to what they would be in a screen-detected population.

Limitations

This section considers limitations of the review methodology. Limitations of the evidence and evidence gaps are discussed in the section above.

This rapid review was conducted in line with the UK NSC requirements for evidence summaries, as described at https://www.gov.uk/government/publications/uk-nscevidencereview-process/appendix-f-requirements-for-uk-nsc-evidence-summaries. All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table 114 (Appendix 6).

Database search terms were restricted by study design (RCTs, non-RCTs and observational studies) using a validated search filter.[52] Further limitations included datelimiting the searches to years where previously conducted SLRs (included in this review) were run. The adaptations of the searches are described in the methods section.

Included publication types

This review only included peer-reviewed journal publications and excluded publications that were not peer-reviewed and grey literature. This may have led to the exclusion of relevant evidence. However, this is an accepted methodological adjustment for a rapid review and is unlikely to miss any pivotal studies.

Language

Only studies published in English were included. There is a possibility that some evidence reported in a language other than English was missed. However, this review was ultimately focusing on evidence relevant to the UK setting, and it could be supposed that publications in non-English languages may be more focused on results applicable to other countries. It is anticipated that this limitation should not exclude any pivotal studies.

Review methodology

Articles were reviewed by a single reviewer in the first instance. A second reviewer examined all included articles, 10% of excluded articles, and any articles where there was uncertainty about inclusion. This is a pragmatic strategy that should have minimised the risk of errors and is an accepted methodological adjustment for a rapid review.

Articles not freely available

Searches for full-text articles were carried out at Cambridge University Library. Any unavailable articles were purchased (unless they were not selected for extraction based on study design or intervention, see the Methods section and below).

Study prioritisation

Due to a sufficiently high number of studies initially included in the review, only studies not of the case-control design (question 2) or RCTs (question 3) were ultimately selected for data extraction. This tiered approach to the study selection process was pre-specified and was utilised so that only the most relevant evidence is initially considered in the review.

Appendix 1 — Search strategy

The search strategy included searches of the databases shown in Table 42. MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase.

Table 42. Summary of electronic database searches and dates

Database	Platform	Searched on date	Date range of search
MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print	Ovid SP	[Date]	1946 to Present
Embase	Ovid SP	[Date]	1974 to 2016 July 01
The Cochrane Library, including: - Cochrane Database of Systematic Reviews (CDSR) - Cochrane Central Register of Controlled Trials (CENTRAL) - Database of Abstracts of Reviews of Effects (DARE)	Wiley Online	[Date]	CDSR: Issue 7 of 12, July 2016

Search Terms

Search terms included combinations of free text and subject headings (Medical Subject Headings [MeSH] for MEDLINE, and Emtree terms for Embase). Search terms for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase are shown in Table 43, and search terms for the Cochrane Library databases are shown in Table 44.

Table 43. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Term Group	#	Search terms	Results
Gestational diabetes and maternal glucose (Q1–3)	1.	exp diabetes, gestational/ or exp pregnancy diabetes mellitus/	44112
	2.	(gestational adj4 diabetes).ti,ab.	35867
	3.	(pregnancy adj4 diabetes).ti,ab.	13518
	4.	gdm.ti,ab.	17565
	5.	(glucose adj4 (pregnan$ or gestation$ or natal or maternal)).ti,ab.	10679
	6.	or/1-5	62918
	7.	exp Hyperglycemia/ or (hyperglycaemia or hyperglycemia).ti,ab.	178094
	8.	((impair$ or reduced) adj2 glucose).ti,ab.	57204
	9.	7 or 8	223490

	10.	exp pregnancy/ or (pregnan$ or gestation$ or prenatal$ or antenatal$ or pre-natal$ or ante-natal$ or maternal$).ti,ab.	2249468
	11.	9 and 10	13820
	12.	6 or 11	68891
Outcomes (Q1)	13.	(macrosomia or large for gestational age or LGA).ti,ab.	15291
	14.	exp fetal macrosomia/ or large for gestational age/	11161
	15.	exp birth injuries/	10560
	16.	((perinatal or labor or labour or birth) adj4 trauma).ti,ab.	4023
	17.	((perinatal or labor or labour or birth) adj4 injur$).ti,ab.	7152
	18.	((perinatal or labor or labour or birth) adj4 complication$1).ti,ab.	14554
	19.	exp obstetric labor complications/	244998
	20.	*dystocia/ or exp shoulder dystocia/	6451
	21.	(shoulder adj4 dystocia).ti,ab.	3597
	22.	(fracture$1 adj4 clavicle$1).ti,ab.	4156
	23.	(fracture$1 adj4 humerus).ti,ab.	10726
	24.	(fracture$1 adj4 shoulder$1).ti,ab.	2391
	25.	(fracture$1 adj4 arm$1).ti,ab.	1296
	26.	"erb$ palsy".ti,ab.	504
	27.	neuropath$.ti,ab.	304043
	28.	exp brachial plexus neuropathies/	5211
	29.	(preeclampsia or pre-eclampsia).ti,ab.	72876
	30.	exp pre-eclampsia/	81363
	31.	(heart adj4 (disorder$1 or disease$1)).ti,ab.	417179
	32.	(cardiovascular adj4 (disorder$1 or disease$1)).ti,ab.	429263
	33.	(cardiac adj4 (disorder$1 or disease$1)).ti,ab.	87900
	34.	exp cardiovascular diseases/	6156362
	35.	exp heart diseases/	2825138
	36.	exp hypoglycemia/	102557
	37.	hypoglyc$.ti,ab.	133244
	38.	exp diabetes mellitus, type 2/	362845
	39.	(("type 2" or "type two" or "type II") adj4 diabet$).ti,ab.	344694
	40.	exp obesity/	685751

	41.	(obesity or obese or bmi or "body mass" or overweight).ti,ab.	1181688
	42.	Intensive care units, neonatal/ or neonatal intensive care unit/ or newborn intensive care/ or (neonatal intensive care unit or ICU or NICU).ti,ab.	219411
	43.	or/13-42	8323289
	44.	exp Infant mortality/	48719
	45.	((neonatal or perinatal or infant) adj2 (mortality or death)).ti,ab.	80923
	46.	or/44-45	107531
	47.	(offspring or son$1 or daughter$1 or child or children or pediatric$1 or paediatric$1).ti,ab.	3293382
	48.	exp "child of impaired parents"/	5414
	49.	exp child/	4331844
	50.	(maternal or mother$2).ti,ab.	851259
	51.	exp mothers/	172974
	52.	or/47-51	5993534
	53.	43 and 52	1027135
	54.	46 and 52	67398
Screening and tests (Q2)	55.	mass screening/ or (screen$ or detect$ or predict$ or identif$ or diagnos$).ti.	3636695
	56.	"sensitivity and specificity"/ or (sensitiv$ or specific$ or accura$ or precis$ or detection rate$ or predictive value$ or likelihood ratio$ or false positive$ or receiver operating characteristic$ or ROC curve$ or AUROC).ti,ab.	10621816
	57.	55 or 56	12922737
	58.	Glucose intolerance/	25699
	59.	Glucose Tolerance Test/	56003
	60.	(glucose adj2 tolerance test).ti,ab.	46234
	61.	(glucose adj2 challenge).ti,ab.	6664
	62.	(IGT or GTT or OGTT or GCT or OGCT).ti,ab.	45778
	63.	(glucose adj3 (test$ or measur$ or assess$ or evaluat$ or monitor$)).ti,ab.	141962
	64.	fasting glucose.ti,ab.	44116
	65.	(maternal history or maternal risk factors or maternal characteristics).ti,ab. or risk assessment/ or risk factors/ or medical history/	2100650
	66.	high risk population/ or high risk pregnancy/ or Pregnancy, High Risk/ or low risk population/ or population risk/	143117
	67.	risk prediction.ti,ab.	20390
	68.	or/58-67	2428313

	69.	Maternal serum screening tests/ or Biomarkers/ or biological marker/ or prenatal diagnosis/	623726
	70.	57 and 68	633071
	71.	57 and 69	287667
Interventions (Q3)			
72. Hypoglycemic Agents/ or antidiabetic agent/ 106527	73.	((pharmacological or hypoglycemic or hypoglycaemic or antihyperglycemic or antihyperglycaemic or antidiabetic or anti-diabetic) adj (agent$ or drug$ or treatment$ or intervention$)).ti,ab.	132937
RCTs (Q1 and Q3)	74.	Metformin/ or Insulin/ or glyburide/ or glybenclamide/ or acarbose/ or sulfonylurea/ or (metformin or insulin or glibenclamide or glimepiride or glipizide or sulfonylurea or sulphonylurea).ti,ab.	985581
	75.	Exercise/ or diet/ or dietary intake/ or food intake/ or maternal nutrition/ or Eating/	938769
	76.	(non-pharmacological or lifestyle modif$ or lifestyle change$ or diet$ or exercis$ or physical activit$).ti,ab.	1964402
	77.	or/72-76	3206482
	78.	exp Randomized Controlled Trials as Topic/	295339
	79.	exp Randomized Controlled Trial/	1054432
	80.	exp Random Allocation/	184039
	81.	exp Randomization/	184039
	82.	exp Double Blind Method/	317306
	83.	exp Single Blind Method/	63463
	84.	exp Single Blind Procedure/	36265
	85.	exp Double Blind Procedure/	164490
	86.	exp Crossover Procedure/	60342
	87.	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).ti,ab,kf.	394322
	88.	exp Clinical Trial/	2255877
	89.	Clinical trial, phase i.pt.	19223
	90.	Clinical trial, phase ii.pt.	31020
	91.	Clinical trial, phase iii.pt.	15408
	92.	Clinical trial, phase iv.pt.	1738
	93.	exp Phase 1 Clinical Trial/ or exp Clinical trial, phase I/	73042
	94.	exp Phase 2 Clinical Trial/ or exp Clinical trial, phase II/	105892
	95.	exp Phase 3 Clinical Trial/ or exp Clinical trial, phase III/	57336
	96.	exp Phase 4 Clinical Trial/ or exp Clinical trial, phase IV/	5292
	97.	Controlled clinical trial.pt.	93227
	98.	Randomized controlled trial.pt.	487724
	99.	Multicenter study.pt.	255180

	100.	Clinical trial.pt.	517538
	101.	exp Clinical Trials as Topic/	633638
	102.	trial$.ti.	641212
	103.	(clinical adj trial$).ti,ab,kf.	831279
	104.	exp Placebos/	374705
	105.	exp Placebo/	340267
	106.	placebo$.ti,ab,kf.	502867
	107.	randomly allocated.ti,ab,kf.	60114
	108.	(allocated adj2 random$).ti,ab,kf.	67093
	109.	random allocation.ti,ab,kf.	3551
	110.	random assignment.ti,ab,kf.	5008
	111.	randomized.ti,ab.	1189654
	112.	randomised.ti,ab.	240885
	113.	randomisation.ti,ab,kf.	20453
	114.	randomization.ti,ab,kf.	67898
	115.	randomly.ti,ab.	736975
	116.	RCT.ti,ab,kf.	53969
	117.	Open-label trial$.ti,ab,kf.	8926
	118.	Open-label stud$.ti,ab,kf.	20621
	119.	Non-blinded stud$.ti,ab,kf.	299
	120.	or/78-119	4530697
Non-RCTs and observational studies (Q1 and Q3)	121.	exp Cohort Studies/	2386872
	122.	exp Cohort Analysis/	2386872
	123.	cohort analy$.ti,ab,kf.	19452
	124.	(cohort adj (study or studies)).ti,ab,kf.	451816
	125.	exp Cross-sectional studies/	613963
	126.	(cross-sectional adj (study or studies)).ti,ab,kf.	342264
	127.	exp Longitudinal Studies/ or exp Longitudinal study/	255242
	128.	Longitudinal.ti,ab,kf.	533505
	129.	exp Follow-Up Studies/	2063152
	130.	exp Follow-Up/	1443120
	131.	(follow up adj (study or studies)).ti,ab,kf.	109521
	132.	exp Prospective Studies/ or exp Prospective study/	1054902

	142.	12 and 53 and (120 or 141)	13625
	143.	12 and 54 and (120 or 141)	1358
	144.	limit 142 to dd=20141001-20190813	6814
	145.	limit 144 to dt=20141001-20190813	4468
	146.	limit 143 to yr=2009-2019	976
	147.	145 or 146	5067
	148.	12 and 70	9126
	149.	12 and 71	1150
	150.	limit 148 to dd=20141001-20190816	4684
	151.	limit 150 to dt=20141001-20190816	2280
	152.	limit 149 to dd=20090101-20190816	896
	153.	limit 152 to dt=20090101-20190816	727
	133.	(Prospective adj (study or studies)).ti,ab,kf.	411385
	134.	(evaluation adj (study or studies)).ti,ab,kf.	12218
	135.	exp Retrospective Studies/ or exp Retrospective study/	1577330
	136.	retrospective$.ti,ab.	1800080
	137.	(chart adj3 review).ti,ab,kf.	111275
	138.	exp Observational studies/ or exp Observational study/	245839
	139.	(observational adj (study or studies)).ti,ab,kf.	246134
	140.	((single arm or single-arm) adj3 (study or studies or trial$)).ti,ab,kf.	14311
	141.	or/121-140	6427392

Question 1

Question 2

	154.	151 or 153	2788
Question 3	155.	12 and 77 and 120	4521
	156.	12 and 77 and 141	6801
	157.	limit 155 to yr=2016-2019	1563
	158.	limit 156 to yr=2009-2019	5204
	159.	157 or 158	6241
Exclusion terms	160.	("Conference Abstract" or "Conference Review" or comment or letter or editorial or note or case reports or news or news release).pt.	9789559
	161.	(case stud$ or case report$).ti,ab.	1004873
	162.	Letter/ or historical article/ or case study/	4307922
	163.	animals/ not humans/	5539852
Combined	165.	147 or 154 or 159	
Remove duplicates	166.	165 not 164	7818
	167.	limit 166 to yr=2016-2019	4750
	168.	166 not 167	3068
	169.	remove duplicates from 167	3437
	170.	remove duplicates from 168	2239
Total	171.	169 or 170	5676
			16130083
	164.	or/160-163	11697

Table 44. Search strategy for the Cochrane Library Databases (Searched via the Wiley Online platform)

Term Group	#	Search terms	Results
Gestational diabetes and maternal glucose (Q1–3)	1.	[mh diabetes, gestational]	785
	2.	(gestational NEAR/4 diabetes):ti,ab	2138
	3.	(pregnancy NEAR/4 diabetes):ti,ab	819

	4.	gdm:ti,ab	1123
	5.	(glucose NEAR/4 (pregnan* or gestation* or natal or maternal)):ti,ab	550
Outcomes (Q1)	6.	{or #1-#5}	2718
	7.	[mh Hyperglycemia] or (hyperglycaemia or hyperglycemia):ti,ab	7005
	8.	((impair* or reduced) NEAR/2 glucose):ti,ab	3324
	9.	#7 or #8	9470
	10.	[mh pregnancy] or (pregnan* or gestation* or prenatal* or antenatal* or prenatal* or ante-natal* or maternal*):ti,ab	66637
	11.	#9 and #10	537
	12.	#6 or #11	2937
	13.	(macrosomia or "large for gestational age" or LGA):ti,ab	651
	14.	[mh "fetal macrosomia"] or [mh ^"large for gestational age"]	111
	15.	[mh "birth injuries"]	42
	16.	((perinatal or labor or labour or birth) NEAR/4 trauma):ti,ab	165
	17.	((perinatal or labor or labour or birth) NEAR/4 injur*):ti,ab	187
	18.	((perinatal or labor or labour or birth) NEAR/4 complication?):ti,ab	747
	19.	[mh "obstetric labor complications"]	3426
	20.	[mh ^dystocia]	97
	21.	(shoulder NEAR/4 dystocia):ti,ab	164
	22.	(fracture? NEAR/4 clavicle?):ti,ab	138
	23.	(fracture? NEAR/4 humerus):ti,ab	334
	24.	(fracture? NEAR/4 shoulder?):ti,ab	126
	25.	(fracture? NEAR/4 arm?):ti,ab	105
	26.	"erb* palsy":ti,ab	5
	27.	neuropath*:ti,ab	10550
	28.	[mh "brachial plexus neuropathies"]	58
	29.	(preeclampsia or pre-eclampsia):ti,ab	2790
	30.	[mh pre-eclampsia]	854
	31.	(heart NEAR/4 (disorder? or disease?)):ti,ab	17480
	32.	(cardiovascular NEAR/4 (disorder? or disease?)):ti,ab	20020
	33.	(cardiac NEAR/4 (disorder? or disease?)):ti,ab	3020
	34.	[mh "cardiovascular diseases"]	97543
	35.	[mh "heart diseases"]	47768
	36.	[mh hypoglycemia]	2014

	37.	hypoglyc*:ti,ab	10813
	38.	[mh "diabetes mellitus, type 2"]	15458
	39.	(("type 2" or "type two" or "type II") NEAR/4 diabet*):ti,ab	32270
	40.	[mh obesity]	12309
	41.	(obesity or obese or bmi or "body mass" or overweight):ti,ab	72399
	42.	[mh ^Intensive care units, neonatal]	667
	43.	("neonatal intensive care unit" or ICU or NICU):ti,ab	13808
	44.	{or #13-#43}	240981
	45.	[mh "Infant mortality"]	629
	46.	((neonatal or perinatal or infant) NEAR/2 (mortality or death)):ti,ab	2533
	47.	#45 or #46	2813
	48.	(offspring or son? or daughter? or child or children or pediatric? or paediatric?):ti,ab	123519
	49.	[mh "child of impaired parents"]	165
	50.	[mh child]	1188
	51.	(maternal or mother*):ti,ab	26943
	52.	[mh mothers]	1602
	53.	{or #48-#52}	141291
	54.	#44 and #53	19115
	55.	#47 and #53	1881
Screening and tests (Q2)	56.	[mh ^"mass screening"] or (screen* or detect* or predict* or identif* or diagnos*):ti	55752
	57.	[mh ^"sensitivity and specificity"] or (sensitiv* or specific* or accura* or precis* or "detection rate*" or "predictive value*" or "likelihood ratio*" or "false positive*" or "receiver operating characteristic*" or "ROC curve*" or AUROC):ti,ab	192058
	58.	#56 or #57	226958
	59.	[mh ^"Glucose intolerance"]	973
	60.	[mh ^"Glucose Tolerance Test"]	1949
	61.	(glucose NEAR/2 "tolerance test"):ti,ab	3094
	62.	(glucose NEAR/2 challenge):ti,ab	322
	63.	(IGT or GTT or OGTT or GCT or OGCT):ti,ab	3003
	64.	(glucose NEAR/3 (test* or measur* or assess* or evaluat* or monitor*)):ti,ab	11639
	65.	"fasting glucose":ti,ab	4058

	66.	("maternal history" or "maternal risk factors" or "maternal characteristics"):ti,ab or [mh ^"risk assessment"] or [mh ^"risk factors"]	29826
	67.	[mh ^"Pregnancy, High-Risk"]	168
	68.	"risk prediction":ti,ab	564
	69.	107-#68	57529
	70.	[mh ^"maternal serum screening tests"] or [mh ^Biomarkers] or [mh ^"prenatal diagnosis"]	13365
	71.	#58 and #69	11921
	72.	#58 and #70	4028
Interventions (Q3)	73.	[mh ^"Hypoglycemic Agents"]	7219
	74.	((pharmacological or hypoglycemic or hypoglycaemic or antihyperglycemic or antihyperglycaemic or antidiabetic or anti-diabetic) NEXT (agent* or drug* or treatment* or intervention*)):ti,ab	9137
	75.	[mh ^Metformin] or [mh ^Insulin] or [mh ^glyburide] or [mh ^acarbose] or (metformin or insulin or glibenclamide or glimepiride or glipizide or sulfonylurea or sulphonylurea):ti,ab	47468
	76.	[mh ^Exercise] or [mh ^diet] or [mh ^eating]	21979
Question 1	77.	(non-pharmacological or "lifestyle modif*" or "lifestyle change*" or diet* or exercis* or "physical activit*"):ti,ab	128444
	78.	{or #73-#77}	174952
	79.	#12 and #54 limit to Cochrane Library publication date from Oct 2014 to Dec 2019, in Cochrane reviews limit to publication year from 2014 to 2019, in Trials	1027 33 669
	80.	#12 and #55 limit to Cochrane Library publication date from Oct 2014 to Dec 2019, in Cochrane reviews limit to publication year from 2014 to 2019, in Trials	108 28 64
	81.	#79 or #80	728
Question 2	82.	#12 and #71 limit to Cochrane Library publication date from Oct 2014 to Dec 2019, in Cochrane reviews limit to publication year from 2014 to 2019, in Trials	350 10 194
	83.	#12 and 72 limit to Cochrane Library publication date from Jan 2009 to Dec 2019, in Cochrane reviews limit to publication year from 2009 to 2019, in Trials	24 0 21
	84.	#82 or #83	217

Question 3	#12 and #78	1755
	85. limit to Cochrane Library publication date from Jan 2016 to Dec 2019, in Cochrane reviews	20
	limit to publication year from 2016 to 2019, in Trials	702
Combined total	#81 or #84 or #85 in Cochrane Reviews in	1241 391202
	86. Trials	

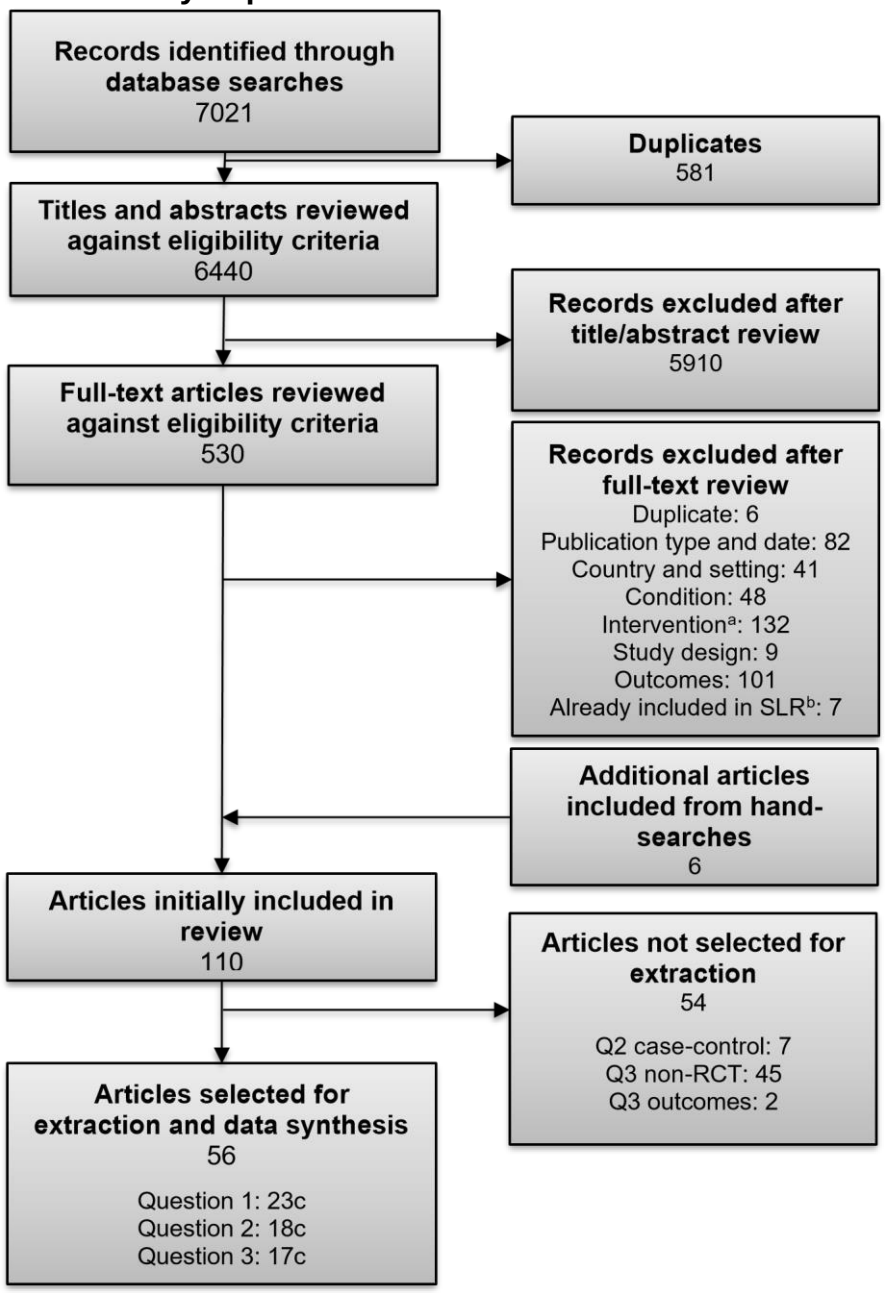
Results were imported into EndNote and de-duplicated.

Appendix 2 — Included and excluded studies

PRISMA flowchart

Figure 1 summarises the volume of publications included and excluded at each stage of the review; 110 publications were ultimately judged to be relevant to 1 or more review questions and were considered for extraction. Publications that were included or excluded after the review of fulltext articles are detailed below

Figure 1. Summary of publications included and excluded at each stage of the review



[a]For Q1: maternal blood glucose not investigated as prognostic factor. For Q2: study not investigating a screening test for GDM. For Q3: study not investigating a pharmacological or lifestyle intervention for GDM. [b]Any records that have already been included in the Farrar 2016 or Brown SLRs that formed the evidence base for this rapid review were excluded. [c]The Farrar 2016 SLR was included for each question, therefore the individual numbers add up to more than the total number of articles included for data extraction.

Publications included after review of full-text articles

110 publications were included after review of full texts. Studies were prioritised for extraction and data synthesis. It was planned *a priori* that the following approach would be taken to prioritise studies for extraction:

1. Systematic reviews and meta-analyses would be considered the highest quality of evidence if any were found.
2. Studies relating to epidemiology would be prioritised if they considered a UK population, followed by studies from Western populations analogous to the UK.

Due to no other SLRs than Farrar 2016 and Brown 2018 L/A/I being included, and few UKspecific studies identified, an *a posteriori* deprioritisation strategy included the following prioritisation by study design:

a. For Question 2, cross-sectional, prospective and retrospective ahead of casecontrol studies
b. For Question 3, RCTs ahead of non-RCTs and observational studies

Publications excluded after review of full-text articles

Of the 530 publications included after the review of titles and abstracts, 420 were ultimately judged not to be relevant to this review. These publications, along with reasons for exclusion, are listed in Table 45.

Table 45. Publications excluded after review of full-text articles

Reference	Reason for exclusion
Abebe KZ, Scifres C, Simhan HN, et al. (2017) Comparison of Two Screening Strategies for Gestational Diabetes (GDM^2 Trial: Design and rationale. Contemporary Clinical Trials 62:43-49	Not a relevant study or publication type
Aceti A, Santhakumaran S, Logan KM, et al. The diabetic pregnancy and offspring blood pressure in childhood: a systematic review and meta-analysis. Diabetologia 2012;55:3114-3127.	Not a relevant intervention
Actrn. Study examining the effects of altering a diet's macronutrient composition on plasma ketone levels in women with gestational diabetes mellitus. Http://www.who.int/trialsearch/trial2.aspx? Trialid=actrn12616000018415 2016.	Published pre-2009
Aker SS, Yuce T, Kalafat E, et al. Association of first trimester serum uric acid levels gestational diabetes mellitus development. Turk Jinekoloji ve Obstetrik Dernegi Dergisi 2016;13:71-74.	Not a relevant intervention
Aksoy H, Aksoy U, Acmaz G, et al. The effect of impaired 50-gram oral glucose challenge test on fetal abdominal wall thickness. Diabetes/metabolism research and reviews 2014;30:570-574.	Not reporting a relevant outcome
Ali MM, Brown M, Karnitis VJ. Third trimester insulin levels are not correlated with fetal macrosomia or delivery complications. Journal of Reproductive Medicine 2014;59:293-298.	Not in a relevant population
Allard C, Sahyouni E, Menard J, et al. Gestational diabetes mellitus identification based on self-monitoring of blood glucose. Canadian Journal of Diabetes 2015;39:162-8.	Not reporting a relevant outcome
Allehdan SS, Basha AS, Asali FF, et al. Dietary and exercise interventions and glycemic control and maternal and newborn outcomes in women diagnosed with gestational diabetes: Systematic review. Diabetes and Metabolic Syndrome: Clinical Research and Reviews 2019;13:2775-2784.	Not a relevant study or publication type
Alptekin H, Cizmecioaylu A, Isik H, et al. Predicting gestational diabetes mellitus during the first trimester using anthropometric measurements and HOMA-IR. Journal of Endocrinological Investigation 2016;39:577-583.	Not a relevant intervention
Alqudah A, McKinley MC, McNally R, et al. Risk of pre-eclampsia in women taking metformin: a systematic review and meta-analysis. Diabetic Medicine 2018;35:160-172.	Not reporting a relevant outcome
Alunni ML, Roeder HA, Moore TR, et al. First trimester gestational diabetes screening - Change in incidence and pharmacotherapy need. Diabetes Research and Clinical Practice 2015;109:135-140.	Not reporting a relevant outcome
Alwan N, Tuffnell Derek J, West J. Treatments for gestational diabetes. Cochrane Database of Systematic Reviews: Reviews 2009;Issue 3.	Published pre-2009
Anand SS, Gupta MK, Schulze KM, et al. What accounts for ethnic differences in newborn skinfold thickness comparing South Asians and White Caucasians? Findings from the START and FAMILY Birth Cohorts. International Journal of Obesity 2016;40:239-44.	Not reporting a relevant outcome
Anand SS, Vasudevan A, Gupta M, Morrison K, Kurpad A, Teo KK, Srinivasan K, Investigators SCS (2013) Rationale and design of South Asian Birth Cohort (START): a Canada-India collaborative study. BMC Public Health 13:79	Not a relevant study or publication type

Anastasiou E, Vasileiou V, Athanasiadou A, et al. Phenotypic and metabolic characteristics of women with isolated hyperglycemia in pregnancy-Is the time-point important? Diabetes Research and Clinical Practice 2010;90:333-338.	Not reporting a relevant outcome
Andrade-Castellanos CA, Colunga-Lozano LE, Delgado-Figueroa N, et al. Subcutaneous rapid-acting insulin analogues for diabetic ketoacidosis. Cochrane Database of Systematic Reviews 2016;2016 (1) (no pagination).	Published pre-2009
Anzolin G, Silva J, Wolff LC, et al. Use of metformin prophylatic in gestacional diabetes mellitus. International journal of gynaecology and obstetrics 2018;143:718-719.	Published pre-2009
Ardilouze A, Bouchard P, Hivert MF, et al. Self-Monitoring of Blood Glucose: A Complementary Method Beyond the Oral Glucose Tolerance Test to Identify Hyperglycemia During Pregnancy. Canadian Journal of Diabetes. 2019.	Not reporting a relevant outcome
Ardilouze JL, Ménard J, Perron P, et al. Gestational diabetes mellitus: the first prospective randomised study of metformine-glyburide vs insulin. Diabetologia 2014;57:S449-S450.	Published pre-2009
Assaf-Balut C, de la Torre NG, Fuentes M, et al. A high adherence to six food targets of the mediterranean diet in the late first trimester is associated with a reduction in the risk of materno-foetal outcomes: The st. carlos gestational diabetes mellitus prevention study. Nutrients 2019;11 (1) (no pagination).	Not a relevant intervention
Assaf-Balut C, Garcia De La Torre N, Duran A, et al. Medical nutrition therapy for gestational diabetes mellitus based on Mediterranean Diet principles: A subanalysis of the St Carlos GDM Prevention Study. BMJ Open Diabetes Research and Care 2018;6 (1) (no pagination).	Not a relevant intervention
Aydin H, Celik O, Yazici D, et al. Prevalence and predictors of gestational diabetes mellitus: a nationwide multicentre prospective study. Diabetic Medicine 2019;36:221-227.	Not reporting a relevant outcome
Balsells M, Garcia-Patterson A, Sola I, et al. Glibenclamide, metformin, and insulin for the treatment of gestational diabetes: a systematic review and meta-analysis. BMJ 2015;350:h102.	Not a relevant study or publication type
Bao H, Yu P, Song X, et al. The influence of home-based exercise on gestational diabetes: a meta-analysis of randomized controlled trials. Journal of Maternal-Fetal & Neonatal Medicine 2019:1-6.	Not reporting a relevant outcome
Barbour LA, Farabi SS, Friedman JE, et al. Postprandial Triglycerides Predict Newborn Fat More Strongly than Glucose in Women with Obesity in Early Pregnancy. Obesity 2018;26:1347-1356.	Not reporting a relevant outcome
Bartels HC, O'Connor C, Segurado R, et al. Fetal growth trajectories and their association with maternal and child characteristics. Journal of maternal-fetal & neonatal medicine 2018.	Not in a relevant population
Benaiges D, Flores-Le Roux JA, Marcelo I, et al. Is first-trimester HbA1c useful in the diagnosis of gestational diabetes? Diabetes Research & Clinical Practice 2017;133:85-91.	Not a relevant intervention
Benhalima K, Van Crombrugge P, Moyson C, et al. Characteristics and pregnancy outcomes across gestational diabetes mellitus subtypes based on insulin resistance. Diabetologia. 2019.	Not reporting a relevant outcome
Berger H, Gagnon R, Sermer M. No. 390-Diabetes in Pregnancy. Journal of Obstetrics and Gynaecology Canada. 2019.	Not a relevant study or publication type
Berger H, Sermer M, Basso M, et al. Diabetes in Pregnancy. Journal of Obstetrics and Gynaecology Canada 2016;38:667-679.e1.	Published pre-2009
Berggren EK, Boggess KA, Mathew L, et al. First trimester maternal glycated hemoglobin and sex hormone-binding globulin do not predict third trimester glucose intolerance of pregnancy. Reproductive Sciences 2017;24:613-618.	Not reporting a relevant outcome
Berggren EK, Roeder HA, Boggess KA, et al. First-trimester maternal serum C-reactive protein as a predictor of third-trimester impaired glucose tolerance. Reproductive Sciences 2015;22:90-93.	Not reporting a relevant outcome

Berglund SK, Garcia-Valdes L, Torres-Espinola FJ, et al. Maternal, fetal and perinatal alterations associated with obesity, overweight and gestational diabetes: an observational cohort study (PREOBE). BMC public health 2016;16:207.	Not a relevant intervention
Bgeginski R, Ribeiro PAB, Mottola MF, et al. Effects of weekly supervised exercise or physical activity counseling on fasting blood glucose in women diagnosed with gestational diabetes mellitus: A systematic review and meta-analysis of randomized trials. Journal of Diabetes 2017;9:1023-1032.	Published pre-2009
Bianchi C, de Gennaro G, Romano M, et al. Italian national guidelines for the screening of gestational diabetes: Time for a critical appraisal? Nutrition, Metabolism and Cardiovascular Diseases 2017;27:717-722.	Not reporting a relevant outcome
Black MH, Sacks DA, Xiang AH, et al. Clinical outcomes of pregnancies complicated by mild gestational diabetes mellitus differ by combinations of abnormal oral glucose tolerance test values. Diabetes Care 2010;33:2524-2530.	Not a relevant intervention
Black MH, Sacks DA, Xiang AH, et al. The relative contribution of prepregnancy overweight and obesity, gestational weight gain, and IADPSG-defined gestational diabetes mellitus to fetal overgrowth. Diabetes Care 2013;36:56-62.	Not a relevant intervention
Blais S, Patenaude J, Doyon M, et al. Effect of gestational diabetes and insulin resistance on offspring's myocardial relaxation kinetics at three years of age. PLoS ONE 2018;13 (11) (no pagination).	Not a relevant intervention

Boisvert MR, Koski KG, Burns DH, et al. Prediction of gestational diabetes mellitus based on an analysis of amniotic fluid by capillary electrophoresis. Biomarkers in Medicine 2012;6:645-53.	Not in a relevant population
Bourdages M, Demers ME, Dube S, et al. First-Trimester Abdominal Adipose Tissue Thickness to Predict Gestational Diabetes. Journal of Obstetrics and Gynaecology Canada 2018;40:883-887.	Not a relevant intervention
Boutzios G, Livadas S, Piperi C, et al. Polycystic ovary syndrome offspring display increased oxidative stress markers comparable to gestational diabetes offspring. Fertility & Sterility 2013;99:943-50.	Not a relevant intervention
Brand JS, West J, Tuffnell D, et al. Gestational diabetes and ultrasound-assessed fetal growth in South Asian and White European women: findings from a prospective pregnancy cohort. BMC Medicine 2018;16:203.	Not reporting a relevant outcome
Brown J, Ceysens G, Boulvain M. Exercise for pregnant women with gestational diabetes for improving maternal and fetal outcomes. Cochrane Database of Systematic Reviews 2017;2017 (6) (no pagination).	Not a relevant study or publication type
Brustman L, Langer O, Scarpelli S, et al. Hypoglycemia in glyburide-treated gestational diabetes: is it dose-dependent? Obstetrics & Gynecology 2011;117:349-53.	Not a relevant intervention
Budak MS, Arac E. Maternal hypoglycaemia on the 50 g oral glucose challenge test - evaluation of obstetric and neonatal outcomes. Ginekologia polska 2018;89:370-374.	Not a relevant intervention
Buitrago-Leal M, Molina-Giraldo S. Safety and effectiveness of oral hypoglycemic agents in women with gestational diabetes. Literature Review. Ginecologia y Obstetricia de Mexico 2014;82:454-464.	Published pre-2009
Burris HH, Rifas-Shiman SL, Kleinman K, et al. Vitamin D deficiency in pregnancy and gestational diabetes mellitus. American Journal of Obstetrics and Gynecology 2012;207:182.e1-182.e8.	Not reporting a relevant outcome
Bush NC, Chandler-Laney PC, Rouse DJ, et al. Higher maternal gestational glucose concentration is associated with lower offspring insulin sensitivity and altered beta-cell function. Journal of Clinical Endocrinology and Metabolism 2011;96:E803-E809.	Not reporting a relevant outcome
Butalia S, Gutierrez L, Lodha A, et al. Short- and long-term outcomes of metformin compared with insulin alone in pregnancy: a systematic review and meta-analysis. Diabetic Medicine 2017;34:27-36.	Not a relevant study or publication type
Caglar GS, Ozdemir EDU, Cengiz SD, et al. Sex-hormone-binding globulin early in pregnancy for the prediction of severe gestational diabetes mellitus and related complications. Journal of Obstetrics and Gynaecology Research 2012;38:1286-1293.	Not a relevant intervention

Caissutti C, Khalifeh A, Saccone G, et al. Are women positive for the One Step but negative for the Two Step screening tests for gestational diabetes at higher risk for adverse outcomes? Acta Obstetricia et Gynecologica Scandinavica 2018;97:122-134.	Not reporting a relevant outcome
Caissutti C, Saccone G, Ciardulli A, et al. Very tight vs. tight control: what should be the criteria for pharmacologic therapy dose adjustment in diabetes in pregnancy? Evidence from randomized controlled trials. Acta Obstetricia et Gynecologica Scandinavica 2018;97:235-247.	Not reporting a relevant outcome
Caissutti C, Saccone G, Khalifeh A, et al. Which criteria should be used for starting pharmacologic therapy for management of gestational diabetes in pregnancy? Evidence from randomized controlled trials. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians 2019;32:2905-2914.	Not a relevant study or publication type
Capula C, Chiefari E, Borelli M, et al. A new predictive tool for the early risk assessment of gestational diabetes mellitus. Primary care diabetes 2016;10:315-23.	Not a relevant intervention
Carolan-Olah M, Duarte-Gardea M, Lechuga J. A systematic review of interventions for Hispanic women with or at risk of Gestational diabetes mellitus (GDM). Sexual and Reproductive Healthcare 2017;13:14-22.	Not a relevant study or publication type
Carroll DG, Kelley KW. Review of metformin and glyburide in the management of gestational diabetes. Pharmacy Practice 2014;12:528.	Published pre-2009
Casey BM, Duryea EL, Abbassi-Ghanavati M, et al. Glyburide in Women With Mild Gestational Diabetes: A Randomized Controlled Trial. Obstetrics & Gynecology 2015;126:303-9.	Not in a relevant population
Casey C, McGinty A, Holmes VA, et al. Maternal vitamin D and markers of glycaemia during pregnancy in the Belfast centre of the Hyperglycaemia and Adverse Pregnancy Outcome study. Diabetic Medicine 2018;35:972-979.	Not reporting a relevant outcome
Ceysens G, Brown J, Boulvain M. Exercise for pregnant women with gestational diabetes for improving maternal and fetal outcomes. Cochrane Database of Systematic Reviews 2016;2016 (5) (no pagination).	Published pre-2009

Chasan-Taber L, Marcus BH, Rosal MC, et al. Estudio Parto: postpartum diabetes prevention program for hispanic women with abnormal glucose tolerance in pregnancy: a randomised controlled trial - study protocol. BMC pregnancy and childbirth 2014;14:100.	Not a relevant study or publication type
Cheung NW, Jiang S, Athayde N. Impact of the IADPSG criteria for gestational diabetes, and of obesity, on pregnancy outcomes. Australian and New Zealand Journal of Obstetrics and Gynaecology 2018;58:553-559.	Not a relevant intervention
Chu L, Retnakaran R, Zinman B, et al. Impact of maternal physical activity and infant feeding practices on infant weight gain and adiposity. International Journal of Endocrinology 2012;2012 (no pagination).	Not reporting a relevant outcome
Claesson R, Ekelund M, Berntorp K. The potential impact of new diagnostic criteria on the frequency of gestational diabetes mellitus in Sweden. Acta Obstetricia et Gynecologica Scandinavica 2013;92:1223-6.	Not reporting a relevant outcome
Clausen TD, Mortensen EL, Schmidt L, et al. Cognitive Function in Adult Offspring of Women with Gestational Diabetes-The Role of Glucose and Other Factors. PLoS ONE 2013;8 (6) (no pagination).	Not reporting a relevant outcome
Contreras ZA, Ritz B, Virk J, et al. Maternal pre-pregnancy and gestational diabetes, obesity, gestational weight gain, and risk of cancer in young children: a population-based study in California. Cancer Causes & Control 2016;27:1273-85.	Not reporting a relevant outcome
Corrado F, Pintaudi B, Di Vieste G, et al. Italian risk factor-based screening for gestational diabetes. Journal of Maternal-Fetal & Neonatal Medicine 2014;27:1445-8.	Not reporting a relevant outcome
Correa PJ, Venegas P, Palmeiro Y, et al. First trimester prediction of gestational diabetes mellitus using plasma biomarkers: a casecontrol study. Journal of Perinatal Medicine 2019;47:161-168.	Not a relevant intervention
Cosson E, Cussac-Pillegand C, Benbara A, et al. The diagnostic and prognostic performance of a selective screening strategy for diabetes mellitus according to ethnicity in Europe. Journal of Clinical Endocrinology & Metabolism 2014;99:996-1005.	Not a relevant intervention gestational

Costi L, Lockwood C, Munn Z, et al. Women's experience of diabetes and diabetes management in pregnancy: a systematic review of qualitative literature. JBI Database of Systematic Reviews and Implementation Reports 2014;12:176-280.	Not a relevant intervention
Costi L. Women's experience of diabetes and diabetes management in pregnancy: a systematic review of qualitative evidence. JBI Library of Systematic Reviews 2012;10:S325-S333.	Not reporting a relevant outcome
Cremona A, O'Gorman C, Cotter A, et al. Effect of exercise modality on markers of insulin sensitivity and blood glucose control in pregnancies complicated with gestational diabetes mellitus: a systematic review. Obesity Science and Practice 2018;4:455-467.	Not reporting a relevant outcome
Crowther CA, Hague WM, Middleton PF, et al. The IDEAL study: investigation of dietary advice and lifestyle for women with borderline gestational diabetes: a randomised controlled trial - study protocol. BMC Pregnancy and Childbirth 2012;12 (no pagination).	Not a relevant intervention
Crume TL, Ogden L, West NA, et al. Association of exposure to diabetes in utero with adiposity and fat distribution in a multiethnic population of youth: The Exploring Perinatal Outcomes among Children (EPOCH) study. Diabetologia 2011;54:87-92.	Not reporting a relevant outcome
Crume TL, Shapiro AL, Brinton JT, et al. Maternal fuels and metabolic measures during pregnancy and neonatal body composition: The healthy start study. Translational Endocrinology and Metabolism 2015;100:1672-1680.	Not a relevant intervention
Cypryk K, Wender-Ozegowska E, Kosinski M, et al. Analysis of maternal factors and pregnancy outcomes in women with gestational diabetes mellitus. Results of the scientific and educational programme "We care for the Mothers". [Polish]. Diabetologia Kliniczna 2014;3:144-156.	Not a relevant intervention
Dalfra MG, Soldato A, Moghetti P, et al. Diabetic pregnancy outcomes in mothers treated with basal insulin lispro protamine suspension or NPH insulin: A multicenter retrospective Italian study. Journal of Maternal-Fetal and Neonatal Medicine 2016;29:1061-1065.	Not in a relevant population
de Seymour JV, Conlon CA, Sulek K, et al. Early pregnancy metabolite profiling discovers a potential biomarker for the subsequent development of gestational diabetes mellitus. Acta Diabetologica 2014;51:887-90.	Not reporting a relevant outcome
de Wit L, Rademaker D, Voormolen DN, et al. SUGAR-DIP trial: oral medication strategy versus insulin for diabetes in pregnancy, study protocol for a multicentre, open-label, non-inferiority, randomised controlled trial. BMJ Open 2019;9:e029808.	Published pre-2009
Delmis J, Pavic M, Ivanisevic M, et al. Body mass index and pregnancy outcome. Gynaecologia et Perinatologia 2015;24:99-105.	Not a relevant intervention
Dennedy MC, Avalos G, O'Reilly MW, et al. The impact of maternal obesity on gestational outcomes. Irish medical journal 2012;105:2325.	Not a relevant intervention
Deveer R, Deveer M, Akbaba E, et al. The effect of diet on pregnancy outcomes among pregnants with abnormal glucose challenge test. European Review for Medical and Pharmacological Sciences 2013;17:1258-1261.	Not in a relevant population
Dhulkotia JS, Ola B, Fraser R, et al. Oral hypoglycemic agents vs insulin in management of gestational diabetes: a systematic review and metaanalysis. American Journal of Obstetrics and Gynecology 2010;203:457.e1-457.e9.	Not a relevant study or publication type
Dhyani V, Mahantashetti N, Ganachari M, et al. Effect of counseling on maternal and neonatal complications and quality of life (WHOQoL BREF) in pregnant women diagnosed with gestational diabetes mellitus. Diabetes 2016;65:A173-.	Published pre-2009
Diaz SO, Barros AS, Goodfellow BJ, et al. Second trimester maternal urine for the diagnosis of trisomy 21 and prediction of poor pregnancy outcomes. Journal of Proteome Research 2013;12:2946-57.	Not reporting a relevant outcome
Diderholm B, Beardsall K, Murgatroyd P, et al. Maternal rates of lipolysis and glucose production in late pregnancy are independently related to foetal weight. Clinical Endocrinology 2017;87:272-278.	Not reporting a relevant outcome
Doder Z, Vanechanos D, Oster M, et al. Insulin glulisine in pregnancy - Experience from clinical trials and post-marketing surveillance. European Endocrinology 2015;11:17-20.	Not in a relevant population

Dogan K, Guraslan H, Cankaya A, et al. Ischemia-Modified Albumin (IMA): A Novel Marker for Preeclampsia Independent of Uterine Artery Notching Identified by Doppler Ultrasound. Hypertension in Pregnancy 2015;34:516-524.	Not a relevant intervention
Donovan BM, Breheny PJ, Robinson JG, et al. Development and validation of a clinical model for preconception and early pregnancy risk prediction of gestational diabetes mellitus in nulliparous women. PLoS ONE 2019;14 (4) (no pagination).	Not a relevant intervention
Donovan L, Hartling L, Muise M, et al. Screening tests for gestational diabetes: a systematic review for the US Preventive Services Task Force. Annals of Internal Medicine 2013;159:115-122.	Not a relevant study or publication type
Dube MC, Girard M, Morisset AS, et al. Evaluation of Fetal Liver Volume by Tridimensional Ultrasound in Women With Gestational Diabetes Mellitus. Journal of Obstetrics and Gynaecology Canada 2011;33:1095-1098.	Not reporting a relevant outcome
Ducarme G, Desroys Du Roure F, Le Thuaut A, et al. Efficacy of maternal and biological parameters at the time of diagnosis of gestational diabetes mellitus in predicting neonatal morbidity. European Journal of Obstetrics and Gynecology and Reproductive Biology 2018;221:113-118.	Not a relevant intervention
Dunne F, Devane D, Newell J, et al. EMERGE: a randomized placebo controlled trial of early metformin in addition to usual care in the reduction of gestational diabetes mellitus effects. Irish journal of medical science 2016;185:S377-S378.	Published pre-2009
Durnwald CP, Kallan MJ, Allison KC, et al. A Randomized Clinical Trial of an Intensive Behavior Education Program in Gestational Diabetes Mellitus Women Designed to Improve Glucose Levels on the 2-Hour Oral Glucose Tolerance Test. American journal of perinatology 2016;33:1145-1151.	Not a relevant intervention
Egan AM, Heerey AM, Carmody L, et al. The changing diagnosis of gestational diabetes mellitus: does anyone miss out? Diabetes Research & Clinical Practice 2014;106:e53-5.	Not reporting a relevant outcome
Ehrlich SF, Hedderson MM, Xu F, et al. Diagnostic thresholds for pregnancy hyperglycemia, maternal weight status and the risk of childhood obesity in a diverse Northern California cohort using health care delivery system data. PLoS ONE [Electronic Resource] 2019;14:e0216897.	Not in a relevant population
Elkind-Hirsch KE, Ogden BW, Darensbourg CJ, et al. Clinical assessment of insulin action during late pregnancy in women at risk for gestational diabetes: Association of maternal glycemia with perinatal outcome. International Journal of Diabetes Mellitus 2010;2:3-9.	Not reporting a relevant outcome
Ellenberg A, Sarvilinna N, Gissler M, et al. New guidelines for screening, diagnosing, and treating gestational diabetes - evaluation of maternal and neonatal outcomes in Finland from 2006 to 2012. Acta Obstetricia et Gynecologica Scandinavica 2017;96:372-381.	Not a relevant intervention
Ellerbrock J, Bohnen J, Lopes van Balen VA, et al. Homeostatic model assessment of beta cell function predicting abnormal oral glucose tolerance testing in pregnancy: a systematic review and meta-analysis. Gynecological Endocrinology 2017;33:911-917.	Not a relevant study or publication type
Ensenauer R, Brandlhuber L, Burgmann M, et al. Obese Nondiabetic Pregnancies and High Maternal Glycated Hemoglobin at Delivery as an Indicator of Offspring and Maternal Postpartum Risks: The Prospective PEACHES Mother-Child Cohort. Clinical Chemistry 2015;61:1381-90.	Not a relevant intervention
Erjavec K, Poljicanin T, Matijevic R. Impact of the Implementation of New WHO Diagnostic Criteria for Gestational Diabetes Mellitus on Prevalence and Perinatal Outcomes: A Population-Based Study. Journal of pregnancy 2016;2016:2670912.	Not reporting a relevant outcome
Euctr IE. A Clinical Trial of Metformin Versus Placebo for the Treatment of High Blood Sugars During Pregnancy. Http://www.who.int/trialsearch/trial2.aspx? Trialid=euctr2016-001644-19-ie 2016.	Published pre-2009
Euctr NL. The SUGAR-DIP trial: oral medication versus insulin for diabetes in pregnancy. Http://www.who.int/trialsearch/trial2.aspx? Trialid=euctr2016-001401-16-nl 2016.	Published pre-2009

Fadiloglu E, Tanacan A, Unal C, et al. Clinical importance of the 75-g glucose tolerance test (GTT) in the prediction of large for gestational age (LGA) fetuses in non-diabetic pregnancies. Journal of Perinatal Medicine 2019;47:534-538.	Not reporting a relevant outcome
Fadl HE, Ostlund IKM, Magnuson AFK, et al. Maternal and neonatal outcomes and time trends of gestational diabetes mellitus in Sweden from 1991 to 2003. Diabetic Medicine 2010;27:436-441.	Not reporting a relevant outcome
Falavigna M, Schmidt MI, Trujillo J, et al. Effectiveness of gestational diabetes treatment: a systematic review with quality of evidence assessment. Diabetes Research and Clinical Practice 2012;98:396-405.	Not a relevant study or publication type
Farina A, Eklund E, Bernabini D, et al. A First-Trimester Biomarker Panel for Predicting the Development of Gestational Diabetes.	Not a relevant intervention Reproductive Sciences 2017;24:954-959.
Farrant MT, Williamson K, Battin M, et al. The use of dextrose/insulin infusions during labour and delivery in women with gestational diabetes mellitus: Is there any point? Australian and New Zealand Journal of Obstetrics and Gynaecology 2017;57:378-380.	Not a relevant intervention
Farrar D, Duley L, Dowswell T, et al. Different strategies for diagnosing gestational diabetes to improve maternal and infant health. Cochrane Database of Systematic Reviews 2017.	Not in a relevant study type
Farrar D, Duley L, Medley N, et al. Different strategies for diagnosing gestational diabetes to improve maternal and infant health. Cochrane Database of Systematic Reviews 2015;1:CD007122.	Not in a relevant study type
Farrar D, Simmonds M, Bryant M, et al. Treatments for gestational diabetes: A systematic review and meta-analysis. BMJ Open 2017;7 (6) (no pagination).	Not a relevant study or publication type
Farrar D, Tuffnell DJ, West J, et al. Continuous subcutaneous insulin infusion versus multiple daily injections of insulin for pregnant women with diabetes. The Cochrane database of systematic reviews 2016:CD005542.	Not in a relevant population
Feig DS, Murphy K, Asztalos E, et al. Metformin in women with type 2 diabetes in pregnancy (MiTy): A multi-center randomized controlled	Not in a relevant population trial. BMC Pregnancy and Childbirth 2016;16 (1) (no pagination).
Ferrara A, Hedderson MM, Albright CL, et al. A pragmatic cluster randomized clinical trial of diabetes prevention strategies for women with gestational diabetes: design and rationale of the Gestational Diabetes' Effects on Moms (GEM) study. BMC pregnancy and childbirth 2014;14:21.	Not reporting a relevant outcome
Ferrara A, Hedderson MM, Brown SD, et al. The Comparative Effectiveness of Diabetes Prevention Strategies to Reduce Postpartum Weight Retention in Women with Gestational Diabetes Mellitus: The Gestational Diabetes' Effects on Moms (GEM) Cluster Randomized Controlled Trial. Diabetes Care 2016;39:65-74.	Not reporting a relevant outcome
Fong A, Serra AE, Gabby L, et al. Use of hemoglobin A1c as an early predictor of gestational diabetes mellitus. American Journal of Obstetrics and Gynecology 2014;211:641.e1-641.e7.	Not a relevant intervention
Fonseca A, Lopes J, Clode N. Glucose intolerance in the third trimester is not predictive of adverse outcomes. International Journal of Gynaecology & Obstetrics 2019;15:15.	Not a relevant intervention
Friis CM, Qvigstad E, Paasche Roland MC, et al. Newborn Body Fat: Associations with Maternal Metabolic State and Placental Size. PLoS ONE 2013;8 (2) (no pagination).	Not reporting a relevant outcome
Froslie KF, Roislien J, Qvigstad E, et al. Shape information in repeated glucose curves during pregnancy provided significant physiological information for neonatal outcomes. PLoS ONE 2014;9 (3) (no pagination).	Not reporting a relevant outcome
Galindo A, Burguillo AG, Azriel S, et al. Outcome of fetuses in women with pregestational diabetes mellitus. Journal of Perinatal Medicine 2006;34:323-331.	Not in a relevant population
Gerome JM, Bucher LKM, Dogbey G. Effects of implementing international association of diabetes and pregnancy study groups gestational diabetes screening on pregnancy outcomes at a small community teaching hospital. Clinical Diabetes 2017;35:84-89.	Not a relevant intervention
Geurtsen ML, van Soest EEL, Voerman E, et al. High maternal early-pregnancy blood glucose levels are associated with altered fetal growth and increased risk of adverse birth outcomes. Diabetologia. 2019.	Not a relevant intervention

Ghio A, Seghieri G, Lencioni C, et al. 1-hour OGTT plasma glucose as a marker of progressive deterioration of insulin secretion and action in pregnant women. International Journal of Endocrinology 2012;2012 (no pagination).	Not reporting a relevant outcome

Giannakou K, Evangelou E, Yiallouros P, et al. Risk factors for gestational diabetes: An umbrella review of meta-analyses of observational studies. PLoS ONE 2019;14 (4) (no pagination).	Not a relevant intervention
Gilbert L, Gross J, Lanzi S, et al. How diet, physical activity and psychosocial well-being interact in women with gestational diabetes mellitus: An integrative review. BMC Pregnancy and Childbirth 2019;19 (1) (no pagination).	Not in a relevant study type
Gillman MW, Oakey H, Baghurst PA, et al. Effect of treatment of gestational diabetes mellitus on obesity in the next generation. Diabetes Care 2010;33:964-8.	Not a relevant intervention
Glover AV, Tita A, Biggio JR, et al. Examining the Starting Dose of Glyburide in Gestational Diabetes. American Journal of Perinatology 2016;33:214-20.	Not a relevant intervention
Gobl CS, Bozkurt L, Tura A, et al. Assessment of glucose regulation in pregnancy after gastric bypass surgery. Diabetologia 2017;60:2504-2513.	Not in a relevant population
Godang K, Froslie KF, Henriksen T, et al. Seasonal variation in maternal and umbilical cord 25(OH) vitamin D and their associations with neonatal adiposity. European Journal of Endocrinology 2014;170:609-617.	Not a relevant intervention
Gomes CP, Torloni MR, Gueuvoghlanian-Silva BY, et al. Cytokine levels in gestational diabetes mellitus: a systematic review of the literature. American Journal of Reproductive Immunology 2013;69:545-557.	Not a relevant intervention
Gonzalez CD, Alvarinas J, Bagnes MFG, et al. Metformin and pregnancy outcomes: Evidence gaps and unanswered questions. Current Clinical Pharmacology 2019;14:54-60.	Published pre-2009
Gozukara YM, Aytan H, Ertunc D, et al. Role of first trimester total testosterone in prediction of subsequent gestational diabetes mellitus. Journal of Obstetrics and Gynaecology Research 2015;41:193-198.	Not a relevant intervention
Graves E, Hill DJ, Evers S, et al. The impact of abnormal glucose tolerance and obesity on fetal growth. Journal of Diabetes Research 2015;2015 (no pagination).	Not a relevant intervention
Gui J, Liu Q, Feng L. Metformin vs insulin in the management of gestational diabetes: a meta-analysis. Plos One 2013;8:e64585.	Not a relevant study or publication type
Guillemette L, Allard C, Lacroix M, et al. Genetics of Glucose regulation in Gestation and Growth (Gen3G): a prospective prebirth cohort of mother-child pairs in Sherbrooke, Canada. BMJ Open 2016;6:e010031.	Not a relevant intervention
Haddow JE, Metzger BE, Lambert-Messerlian G, et al. Maternal BMI, Peripheral Deiodinase Activity, and Plasma Glucose: Relationships Between White Women in the HAPO Study. The Journal of clinical endocrinology and metabolism 2019;104:2593-2600.	Not reporting a relevant outcome
Hancerliogullari N, Celik HK, Karakaya BK, et al. Effect of Prolonged Fasting Duration on 50 Gram Oral Glucose Challenge Test in the Diagnosis of Gestational Diabetes Mellitus. Hormone & Metabolic Research 2018;50:671-674.	Not a relevant intervention
Harrison AL, Shields N, Taylor NF, et al. Exercise improves glycaemic control in women diagnosed with gestational diabetes mellitus: a systematic review. Journal of physiotherapy 2016;62:188-196.	Not a relevant study or publication type
Harrison S, Petkovic G, Chevassut A, et al. Oxfordshire Women and Their Children's Health (OxWATCH): protocol for a prospective cohort feasibility study. BMJ Open 2015;5:e009282.	Not a relevant intervention
Hartling L, Dryden DM, Guthrie A, et al. Benefits and harms of treating gestational diabetes mellitus: A systematic review and metaanalysis for the U.S. preventive services task force and the national institutes of health office of medical applications of research. Annals of Internal Medicine 2013;159:123-129.	Not reporting a relevant outcome

Hartling L, Dryden DM, Guthrie A, et al. Diagnostic thresholds for gestational diabetes and their impact on pregnancy outcomes: a systematic review. Diabetic Medicine 2014;31:319-331.	Not a relevant study or publication type
Harville EW, Viikari JS, Raitakari OT. Preconception cardiovascular risk factors and pregnancy outcome. Epidemiology 2011;22:724-30.	Not a relevant intervention
Hauta-alus HH, Viljakainen HT, Holmlund-Suila EM, et al. Maternal vitamin D status, gestational diabetes and infant birth size. BMC Pregnancy and Childbirth 2017;17 (1) (no pagination).	Not reporting a relevant outcome
Hayashi A, Oguchi H, Kozawa Y, et al. Daily walking is effective for the management of pregnant women with gestational diabetes mellitus. Journal of Obstetrics and Gynaecology Research 2018;44:1731-1738.	Not reporting a relevant outcome
He XJ, Qin FY, Hu CL, et al. Is gestational diabetes mellitus an independent risk factor for macrosomia: a meta-analysis? Archives of Gynecology and Obstetrics 2015;291:729-735.	Not a relevant study or publication type
Hedderson MM, Xu F, Sridhar SB, et al. A cohort study of maternal cardiometabolic risk factors and primary cesarean delivery in an integrated health system. PLoS ONE [Electronic Resource] 2018;13:e0199932.	Not reporting a relevant outcome
Heerwagen MJR, Gumina DL, Hernandez TL, et al. Placental lipoprotein lipase activity is positively associated with newborn adiposity. Placenta 2018;64:53-60.	Not a relevant intervention
Hefele JG, Santos P, Ritter G, et al. Risk Factors for Shoulder Dystocia: the Impact of Mother's Race and Ethnicity. Journal of Racial & Ethnic Health Disparities 2018;5:333-341.	Not a relevant intervention
Helle EIT, Biegley P, Knowles JW, et al. First Trimester Plasma Glucose Values in Women without Diabetes are Associated with Risk for Congenital Heart Disease in Offspring. Journal of Pediatrics 2018;195:275-278.	Not reporting a relevant outcome
Herrera K, Brustman L, Foroutan J, et al. The importance of fasting blood glucose in screening for gestational diabetes. Journal of MaternalFetal and Neonatal Medicine 2015;28:825-828.	Not in a relevant population
Herrera K, Rosenn B, Foroutan J, et al. A randomized controlled trial of insulin detemir versus insulin NPH for the treatment of pregnant women with gestational diabetes and type 2 diabetes. American journal of obstetrics and gynecology 2015;212:S320-.	Published pre-2009
Herrera K, Rosenn B, Foroutan J, et al. Are perinatal outcomes different in pregnant women treated with insulin detemir versus NPH? Reproductive sciences. 2015;22:250A.	Published pre-2009
Herrera K, Rosenn B, Foroutan J, et al. Insulin detemir vs. NPH: association with maternal weight gain in pregnancy. Diabetes. 2015;64:A675.	Published pre-2009
Herrera KM, Rosenn BM, Foroutan J, et al. Randomized controlled trial of insulin detemir versus NPH for the treatment of pregnant women with diabetes. American journal of obstetrics and gynecology 2015;213:426.e1-7.	Not in a relevant population
Hieronimus S, Le Meaux JP. Relevance of gestational diabetes mellitus screening and comparison of selective with universal strategies. Diabetes & Metabolism 2010;36:575-86.	Published pre-2009
Hiersch L, Berger H, Okby R, et al. Gestational diabetes mellitus is associated with adverse outcomes in twin pregnancies. American Journal of Obstetrics & Gynecology 2019;220:102.e1-102.e8.	Not reporting a relevant outcome
Hill AJ, Cairnduff V, McCance DR. Nutritional and clinical associations of food cravings in pregnancy. Journal of human nutrition and dietetics : the official journal of the British Dietetic Association 2016;29:281-289.	Not a relevant intervention
Hillier TA, Pedula KL, Vesco KK, et al. Impact of Maternal Glucose and Gestational Weight Gain on Child Obesity over the First Decade of Life in Normal Birth Weight Infants. Maternal and child health journal 2016;20:1559-1568.	Not reporting a relevant outcome
Horie I, Kawasaki E, Sakanaka A, et al. Efficacy of nutrition therapy for glucose intolerance in Japanese women diagnosed with gestational diabetes based on IADPSG criteria during early gestation. Diabetes Research and Clinical Practice 2015;107:400-406.	Not reporting a relevant outcome

Horsch A, Gilbert L, Lanzi S, et al. Improving cardiometabolic and mental health in women with gestational diabetes mellitus and their offspring: Study protocol for MySweetHeart Trial, a randomised controlled trial. BMJ Open 2018;8 (2) (no pagination).	Not a relevant intervention
Horsch A, Gilbert L, Lanzi S, et al. Prospective associations between maternal stress during pregnancy and fasting glucose with obstetric and neonatal outcomes. Journal of Psychosomatic Research 2019;125 (no pagination).	Not a relevant intervention
Horvath K, Koch K, Jeitler K, et al. Effects of treatment in women with gestational diabetes mellitus: systematic review and metaanalysis. Bmj 2010;340:c1395.	Not reporting a relevant outcome
Hosseini E, Janghorbani M. Systematic review and meta-analysis of diagnosing gestational diabetes mellitus with one-step or two-step approaches and associations with adverse pregnancy outcomes. International Journal of Gynaecology & Obstetrics 2018;143:137-144.	Not a relevant intervention
Hu Z, Tylavsky FA, Han JC, et al. Maternal metabolic factors during pregnancy predict early childhood growth trajectories and obesity risk: the CANDLE Study. International Journal of Obesity 2019;31:31.	Not a relevant intervention
Hughes RCE, Moore MP, Gullam JE, et al. An early pregnancy HbA<inf>1c</inf> >=5.9% (41 mmol/mol) is optimal for detecting diabetes and identifies women at increased risk of adverse pregnancy outcomes. Diabetes Care 2014;37:2953-2959.	Not a relevant intervention
Huhn EA, Fischer T, Gobl CS, et al. Screening of gestational diabetes mellitus in early pregnancy by oral glucose tolerance test and glycosylated fibronectin: Study protocol for an international, prospective, multicentre cohort trial. BMJ Open 2016;6 (10) (no pagination).	Not reporting a relevant outcome
Huhn EA, Massaro N, Streckeisen S, et al. Fourfold increase in prevalence of gestational diabetes mellitus after adoption of the new International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria. Journal of Perinatal Medicine 2017;45:359-366.	Not reporting a relevant outcome

Huhtala M, Tertti K, Ronnemaa T. Inflammation markers and insulin like growth factor binding protein 1 in women with gestational diabetes treated with metformin or insulin. Diabetologia 2018;61:S461-S462.	Published pre-2009
Hummel S, Much D, Rossbauer M, et al. Postpartum outcomes in women with gestational diabetes and their offspring: POGO study design and first-year results. The Review of Diabetic Studies 2013;10:49-57.	Not reporting a relevant outcome
Ijas H, Vaarasmaki M, Morin-Papunen L, et al. Metformin should be considered in the treatment of gestational diabetes: A prospective randomised study. BJOG: An International Journal of Obstetrics and Gynaecology 2011;118:880-885.	Not in a relevant population
Ijas H, Vaarasmaki M, Saarela T, et al. A follow-up of a randomised study of metformin and insulin in gestational diabetes mellitus: growth and development of the children at the age of 18 months. BJOG: An International Journal of Obstetrics & Gynaecology 2015;122:994-1000.	Not in a relevant population
Ikenoue S, Miyakoshi K, Saisho Y, et al. Clinical impact of women with gestational diabetes mellitus by the new consensus criteria: Two year experience in a single institution in Japan. Endocrine Journal 2014;61:353-358.	Not a relevant intervention
Iliodromiti S, Sassarini J, Kelsey TW, et al. Accuracy of circulating adiponectin for predicting gestational diabetes: a systematic review and meta-analysis. Diabetologia 2016;59:692-9.	Not a relevant intervention
Immanuel J, Simmons D. Screening and Treatment for Early-Onset Gestational Diabetes Mellitus: a Systematic Review and Metaanalysis. Current Diabetes Reports 2017;17 (11) (no pagination).	Not a relevant intervention
Inan C, Varol FG, Erzincan SG, et al. Use of prokineticin-1 (PROK1), pregnancy-associated plasma protein A (PAPP-A) and PROK1/PAPP-A ratio to predict adverse pregnancy outcomes in the first trimester: a prospective study. Journal of Maternal-Fetal & Neonatal Medicine 2018;31:2685-2692.	Not a relevant intervention
Irct2013102315045N. Glibenclamide effect on maternal and neonatal outcomes in women with gestational diabetes. Http://www.who.int/trialsearch/trial2.aspx? Trialid=irct2013102315045n2 2014.	Not a relevant study or publication type
Irct2014010116025N. Metformin is more effective than insulin in treating gestational diabetes. Http://www.who.int/trialsearch/trial2.aspx?	Not a relevant study or publication

Trialid=irct2014010116025n1 2014.	type
Irct2016041317827N. Efficacy of Glibanclamide on Gestational diabetes and outcome of pregnancy. Http://www.who.int/trialsearch/trial2.aspx? Trialid=irct2016041317827n1 2016.	Not a relevant study or publication type
Irct2016080229172N. Comparing the effect of metformin versus insulin in controlling the blood sugar in of gestational diabetic mother. Http://www.who.int/trialsearch/trial2.aspx? Trialid=irct2016080229172n1 2017.	Not a relevant study or publication type
Irct2016112027263N. The Effect of Aqua Training on Visfatin and Glucose Tolerance. Http://www.who.int/trialsearch/trial2.aspx? Trialid=irct2016112027263n3 2016.	Not a relevant study or publication type
Irct20180103038196N. Comparison of the effect of two regimens of metformin and insulin on results of postpartum glucose tolerance test. Http://www.who.int/trialsearch/trial2.aspx? Trialid=irct20180103038196n5 2018.	Not a relevant study or publication type
Isrctn. Impact on pregnancy outcomes when changing diagnostic criteria for gestational diabetes in Sweden. Http://www.who.int/trialsearch/trial2.aspx? Trialid=isrctn41918550 2017.	Not reporting a relevant outcome
Isrctn. Is there a difference between 2 doses of 12 mg of dexamethasone given 12 hours apart compared with 4 doses of 6 mg dexamethasone given 12 hours apart in terms of the dosing regimen's impact on causing high blood sugar in the mother when dexamethasone is given to minimise risk of prematurity complications? Http://www.who.int/trialsearch/trial2.aspx? Trialid=isrctn16613220 2018.	Not a relevant study or publication type
Isrctn. Metformin in mild gestational diabetes mellitus: a double-blind placebo-controlled randomised trial. Http://www.who.int/trialsearch/trial2.aspx? Trialid=isrctn10845466 2017.	Not a relevant study or publication type
Isrctn. Preventing type 2 diabetes with metformin. Http://www.who.int/trialsearch/trial2.aspx? Trialid=isrctn20930880 2018.	Not a relevant intervention
Iwama N, Sugiyama T, Metoki H, et al. Maternal body mass index is a better indicator of large-for-gestational-age infants compared with a 75-g oral glucose tolerance test in early pregnancy: The JAGS trial. Diabetes Research & Clinical Practice 2017;132:10-18.	Not a relevant intervention
Jarvelin MR, Vaarasmaki M, Pouta A, et al. Adolescent manifestations of metabolic syndrome among children born to women with gestational diabetes in a general-population birth cohort. American Journal of Epidemiology 2009;169:1209-1215.	Not reporting a relevant outcome
Jeelani H, Jabeen F, Mushtaq S. Perinatal morbidity and mortality in gestational diabetes mellitus. JK practitioner 2015;20:30-34.	Not a relevant study or publication type
Jenum AK, Richardsen KR, Berntsen S, et al. Gestational diabetes, insulin resistance and physical activity in pregnancy in a multi-ethnic population - a public health perspective. Norsk Epidemiologi 2013;23:45-54.	Not a relevant intervention
Jenum AK, Sletner L, Voldner N, et al. The STORK Groruddalen research programme: A population-based cohort study of gestational diabetes, physical activity, and obesity in pregnancy in a multiethnic population. Rationale, methods, study population, and participation rates. Scandinavian journal of public health 2010;38:60-70.	Not reporting a relevant outcome
Jiang YF, Chen XY, Ding T, et al. Comparative efficacy and safety of OADs in management of GDM: network meta-analysis of randomized controlled trials. Journal of Clinical Endocrinology & Metabolism 2015;100:2071-80.	Not a relevant study or publication type
Jprn U. Prevention study of diabetes mellitus by intensive lifestyle modification in gestational diabetes mothers: daisy M Study. Http://www.who.int/trialsearch/trial2.aspx? Trialid=jprn-umin000027029 2017.	Published pre-2009
Kadakia R, Nodzenski M, Talbot O, et al. Maternal metabolites during pregnancy are associated with newborn outcomes and hyperinsulinaemia across ancestries. Diabetologia 2019;62:473-484.	Not a relevant intervention

Kalafat E, Sukur YE, Abdi A, et al. Metformin for prevention of hypertensive disorders of pregnancy in women with gestational diabetes or obesity: systematic review and meta-analysis of randomized trials. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2018;52:706-714.	Published pre-2009
Kapaya H, May L, Jacques R, et al. Sensitivity of postnatal fasting plasma glucose in identifying impaired glucose tolerance in women with gestational diabetes - 25 Years' data. Obstetric Medicine 2017;10:125-131.	Not a relevant intervention
Kasap E, Genc M, Turan GA, et al. The association of impaired gestational glucose tolerance with maternal and fetal outcomes. Clinical and Experimental Obstetrics and Gynecology 2016;43:755-758.	Not a relevant intervention
Kayemba-Kay's S, Peters C, Geary MPP, et al. Maternal hyperinsulinism and glycaemic status in the first trimester of pregnancy are associated with the development of pregnancy-induced hypertension and gestational diabetes. European Journal of Endocrinology 2013;168:413-418.	Not a relevant intervention
Kedia R, Kulkarni S, Ross M, et al. Spotlight on empagliflozin/metformin fixed-dose combination for the treatment of type 2 diabetes: A systematic review. Patient Preference and Adherence 2016;10:1999-2006.	Not a relevant study or publication type
Kew S, Ye C, Sermer M, et al. Postpartum metabolic function in women delivering a macrosomic infant in the absence of gestational diabetes mellitus. Diabetes Care 2011;34:2608-2613.	Not a relevant intervention
Kgosidialwa O, Egan AM, Carmody L, et al. Treatment with diet and exercise for women with gestational diabetes mellitus diagnosed using IADPSG criteria. Translational Endocrinology and Metabolism 2015;100:4629-4636.	Not a relevant intervention
Kino T, Yamamoto Y, Saigusa Y, et al. Adverse pregnancy outcomes related to preterm cesarean delivery. European Journal of Obstetrics, Gynecology, & Reproductive Biology 2019;234:89-91.	Not in a relevant population
Kokanali MK, Tokmak A, Kaymak O, et al. The effect of treatment on pregnancy outcomes in women with one elevated oral glucose tolerance test value. Ginekologia Polska 2014;85:748-53.	Not a relevant intervention
Kong JM, Lim K, Thompson DM. Evaluation of the International Association of the Diabetes In Pregnancy Study Group new criteria: gestational diabetes project. Canadian Journal of Diabetes 2015;39:128-32.	Not reporting a relevant outcome
Konig AB, Junginger S, Reusch J, et al. Gestational diabetes outcome in a single center study: Higher BMI in children after six months. Hormone and Metabolic Research 2014;46:804-809.	Not a relevant intervention
Koning SH, van Zanden JJ, Hoogenberg K, et al. New diagnostic criteria for gestational diabetes mellitus and their impact on the number of diagnoses and pregnancy outcomes. Diabetologia 2018;61:800-809.	Not in a relevant population
Kramer CK, Swaminathan B, Hanley AJ, et al. Each degree of glucose intolerance in pregnancy predicts distinct trajectories of beta-cell function, insulin sensitivity, and glycemia in the first 3 years postpartum. Diabetes Care 2014;37:3262-3269.	Not reporting a relevant outcome
Laafira A, White SW, Griffin CJ, et al. Impact of the new IADPSG gestational diabetes diagnostic criteria on pregnancy outcomes in Western Australia. Australian & New Zealand Journal of Obstetrics & Gynaecology 2016;56:36-41.	Not a relevant intervention
Lamain-de Ruiter M, Kwee A, Naaktgeboren CA, et al. External validation of prognostic models to predict risk of gestational diabetes mellitus in one Dutch cohort: prospective multicentre cohort study. BMJ (Clinical research ed.) 2016;354:i4338.	Not a relevant intervention
Landon MB, Mele L, Varner MW, et al. The relationship of maternal glycemia to childhood obesity and metabolic dysfunction⁺⁺. Journal of Maternal Fetal and Neonatal Medicine. 2018.	Not reporting a relevant outcome
Lauring JR, Kunselman AR, Pauli JM, et al. Comparison of healthcare utilization and outcomes by gestational diabetes diagnostic criteria. Journal of Perinatal Medicine 2018;46:401-409.	Not reporting a relevant outcome
Lautatzis ME, Goulis DG, Vrontakis M. Efficacy and safety of metformin during pregnancy in women with gestational diabetes mellitus or polycystic ovary syndrome: A systematic review. Metabolism: Clinical and Experimental 2013;62:1522-1534.	Published pre-2009

Lee GT, Satyan MT, Grothusen JD, et al. A retrospective study comparing outcomes in a midwestern US population after introduction of IADPSG guidelines for gestational diabetes. Journal of Maternal-Fetal & Neonatal Medicine 2019;32:67-72.	Not in a relevant population
Legardeur H, Girard G, Journy N, et al. Factors predictive of macrosomia in pregnancies with a positive oral glucose challenge test: Importance of fasting plasma glucose. Diabetes and Metabolism 2014;40:43-48.	Not in a relevant population
Lekva T, Godang K, Michelsen AE, et al. Prediction of Gestational Diabetes Mellitus and Pre-diabetes 5 Years Postpartum using 75 g Oral Glucose Tolerance Test at 14-16 Weeks' Gestation. Scientific Reports 2018;8:13392.	Not a relevant intervention
Lencioni C, Resi V, Romero F, et al. Glucagon-like peptide-1 secretion in women with gestational diabetes mellitus during and after pregnancy. Journal of Endocrinological Investigation 2011;34:e287-e290.	Not a relevant intervention
Lepercq J, Lin J, Hall GC, et al. Meta-analysis of maternal and neonatal outcomes associated with the use of insulin glargine versus NPH insulin during pregnancy. Obstetrics and Gynecology International 2012:Article Number 649070.	Not a relevant study or publication type
Li G, Zhao S, Cui S, et al. Effect comparison of metformin with insulin treatment for gestational diabetes: a meta-analysis based on RCTs. Archives of Gynecology & Obstetrics 2015;292:111-20.	Not a relevant study or publication type
Liang HL, Ma SJ, Xiao YN, et al. Comparative efficacy and safety of oral antidiabetic drugs and insulin in treating gestational diabetes mellitus: An updated PRISMA-compliant network meta-analysis. Medicine (United States) 2017;96 (38) (no pagination).	Not a relevant study or publication type
Linder K, Schleger F, Kiefer-Schmidt I, et al. Gestational diabetes impairs human fetal postprandial brain activity. Translational Endocrinology and Metabolism 2015;100:4029-4036.	Not a relevant intervention
Lindsay KL, Brennan L, Kennelly MA, et al. Maternal metabolic response to dietary treatment for impaired glucose tolerance and gestational diabetes mellitus. Irish Journal of Medical Science 2018;187:701-708.	Not in a relevant population
Li-Zhen L, Yun X, Xiao-Dong Z, et al. Evaluation of guidelines on the screening and diagnosis of gestational diabetes mellitus: systematic review. BMJ Open 2019;9:e023014.	Published pre-2009
Long DS, Hou W, Taylor RS, et al. Serum levels of endothelial glycocalyx constituents in women at 20 weeks' gestation who later develop gestational diabetes mellitus compared to matched controls: a pilot study. BMJ Open 2016;6:e011244.	Not reporting a relevant outcome
Lowe WL, Bain JR, Nodzenski M, et al. Maternal BMI and glycemia impact the fetal metabolome. Diabetes Care 2017;40:902-910.	Not in a relevant population
Lu L, Koulman A, Petry CJ, et al. An Unbiased Lipidomics Approach Identifies Early Second Trimester Lipids Predictive of Maternal Glycemic Traits and Gestational Diabetes Mellitus. Diabetes Care 2016;39:2232-2239.	Not a relevant intervention
Luo ZC, Delvin E, Fraser WD, et al. Maternal glucose tolerance in pregnancy affects fetal insulin sensitivity. Diabetes Care 2010;33:20552061.	Not reporting a relevant outcome
Maayan-Metzger A, Schushan-Eisen I, Strauss T, et al. Gestational weight gain and body mass indexes have an impact on the outcomes of diabetic mothers and infants. Acta Paediatrica, International Journal of Paediatrics 2015;104:1150-1155.	Not reporting a relevant outcome
Mane L, Flores-Le Roux JA, Benaiges D, et al. Role of First-Trimester HbA1c as a Predictor of Adverse Obstetric Outcomes in a Multiethnic Cohort. Journal of Clinical Endocrinology & Metabolism 2017;102:390-397.	Not a relevant intervention
Mane L, Flores-Le Roux JA, Pedro-Botet J, et al. Is fasting plasma glucose in early pregnancy a better predictor of adverse obstetric outcomes than glycated haemoglobin? European Journal of Obstetrics, Gynecology, & Reproductive Biology 2019;234:79-84.	Not a relevant intervention
Mansell E, Lunt H, Docherty P. Laboratory diagnosis of gestational diabetes: An in silico investigation into the effects of pre-analytical processing on the diagnostic sensitivity and specificity of the oral glucose tolerance test. Clinical Biochemistry 2017;50:506-512.	Not a relevant study or publication type
Martis R, Brown J, Alsweiler J, et al. Different intensities of glycaemic control for women with gestational diabetes mellitus. Cochrane Database of Systematic Reviews 2016.	Published pre-2009

Martis R, Brown J, Alsweiler J, et al. Treatments for women with gestational diabetes mellitus: An overview of Cochrane systematic reviews. Cochrane Database of Systematic Reviews 2016;2016 (8) (no pagination).	Published pre-2009
Martis R, Brown J, Alsweiler J, et al. What targets are best for glycaemic control in women with gestational diabetes mellitus? A cochrane systematic review. Journal of paediatrics and child health 2016;52:42-.	Published pre-2009
Martis R, Crowther CA, Shepherd E, et al. Treatments for women with gestational diabetes mellitus: an overview of Cochrane systematic reviews. Cochrane Database of Systematic Reviews 2018.	Published pre-2009
Matsumoto Y, Yamada H, Yoshida M, et al. Background Factors Determining the Introduction and Dosage of Insulin in Women With Gestational Diabetes Mellitus. Journal of Clinical Medicine Research 2019;11:447-451.	Not reporting a relevant outcome
Matyas E, Horvath K, Jeitler K, et al. Short acting insulin analogues versus regular human insulin in women with gestational diabetes - patient-relevant outcomes: systematic review. Diabetologia 2009;52:S459-S460.	Published pre-2009
McIntyre HD, Chang AM, Callaway LK, et al. Hormonal and metabolic factors associated with variations in insulin sensitivity in human pregnancy. Diabetes Care 2010;33:356-360.	Not a relevant intervention
McIntyre HD, Gibbons KS, Lowe J, Oats JJN (2018) Reprint of "Development of a risk engine relating maternal glycemia and body mass index to pregnancy outcomes". Diabetes Res Clin Pract 145:31-38	Not reporting a relevant outcome
McLeod DSA, Warner JV, Henman M, et al. Associations of serum vitaminD concentrations with obstetric glucose metabolism in a subset of the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study cohort. Diabetic Medicine 2012;29:e199-e204.	Not a relevant intervention
Meek CL, Murphy HR, Simmons D. Random plasma glucose in early pregnancy is a better predictor of gestational diabetes diagnosis than maternal obesity. Diabetologia 2016;59:445-52.	Not a relevant intervention
Meertens LJE, Scheepers HC, De Vries RG, et al. External Validation Study of First Trimester Obstetric Prediction Models (Expect Study I): Research Protocol and Population Characteristics. JMIR Research Protocols 2017;6:e203.	Not in a relevant population
Mendez-Figueroa H, Maggio L, Dahlke JD, et al. Treatment of severe insulin resistance in pregnancy with 500 units per milliliter of concentrated insulin. Obstetrics & Gynecology 2013;122:99-104.	Not in a relevant population
Mert M, Purcu S, Soyluk O, et al. The relationship between glycated hemoglobin and blood glucose levels of 75 and 100 gram oral glucose tolerance test during gestational diabetes diagnosis. International journal of clinical and experimental medicine 2015;8:1333540.	Not in a relevant population
Metcalfe A, Sabr Y, Hutcheon JA, et al. Trends in obstetric intervention and pregnancy outcomes of Canadian women with diabetes in pregnancy from 2004 to 2015. Journal of the Endocrine Society 2018;1:1540-1549.	Not a relevant intervention
Meyer DM, Brei C, Stecher L, et al. Maternal insulin resistance, triglycerides and cord blood insulin are not determinants of offspring growth and adiposity up to 5 years: a follow-up study. Diabetic Medicine 2018;35:1399-1403.	Not a relevant intervention
Miailhe G, Kayem G, Girard G, et al. Selective rather than universal screening for gestational diabetes mellitus? European Journal of Obstetrics Gynecology and Reproductive Biology 2015;191:95-100.	Not reporting a relevant outcome
Ming WK, Ding W, Zhang CJP, et al. The effect of exercise during pregnancy on gestational diabetes mellitus in normal-weight women: A systematic review and meta-analysis. BMC Pregnancy and Childbirth 2018;18 (1) (no pagination).	Not in a relevant population
Mission JF, Catov J, Deihl TE, et al. Early Pregnancy Diabetes Screening and Diagnosis: Prevalence, Rates of Abnormal Test Results, and Associated Factors. Obstetrics & Gynecology 2017;130:1136-1142.	Not in a relevant population
Mitanchez D, Jacqueminet S, Nizard J, et al. Effect of maternal obesity on birthweight and neonatal fat mass: A prospective clinical trial. PLoS ONE 2017;12 (7) (no pagination).	Not a relevant intervention
Monroy G, Tundidor D, Orellana I, et al. Antenatal oral glucose tolerance test in women with gestational diabetes mellitus: fasting plasma glucose is the best predictor of both large for-gestational-age newborns and postpartum glucose tolerance. Minerva Endocrinologica 2017;42:311-317.	Not in a relevant population

Morisset AS, Cote JA, Michaud A, et al. Dietary intakes in the nutritional management of gestational diabetes mellitus. Canadian journal of dietetic practice and research : a publication of Dietitians of Canada = Revue canadienne de la pratique et de la recherche en dietetique : une publication des Dietetistes du Canada 2014;75:64-71.	Not a relevant intervention
Mortier I, Blanc J, Tosello B, et al. Is gestational diabetes an independent risk factor of neonatal severe respiratory distress syndrome after 34 weeks of gestation? A prospective study. Archives of Gynecology & Obstetrics 2017;296:1071-1077.	Not a relevant intervention
Morton S, Kirkwood S, Thangaratinam S. Interventions to modify the progression to type 2 diabetes mellitus in women with gestational diabetes: a systematic review of literature. Current Opinion in Obstetrics and Gynecology 2014;26:476-486.	Published pre-2009
Mosimann B, Amylidi S, Risch L, et al. First-Trimester Placental Growth Factor in Screening for Gestational Diabetes. Fetal Diagnosis & Therapy 2016;39:287-91.	Not in a relevant population
Mutlu N, Esra H, Begum A, et al. Relation of maternal vitamin D status with gestational diabetes mellitus and perinatal outcome. African Health Sciences 2015;15:523-531.	Not a relevant intervention
Nachum Z, Yefet E. Glyburide versus metformin and their combination for the treatment of gestational diabetes mellitus: a randomized controlled study. Diabetes care 2017;40:e116.	Not a relevant intervention
Nagalla SR, Snyder CK, Michaels JE, et al. Maternal serum biomarkers for risk assessment in gestational diabetes. A potential universal screening test to predict GDM status. Indian Journal of Endocrinology and Metabolism 2015;19:155-9.	Published pre-2009
Nansseu JRN, Fokom-Domgue J, Noubiap JJN, et al. Fructosamine measurement for diabetes mellitus diagnosis and monitoring: A systematic review and meta-analysis protocol. BMJ Open 2015;5 (5) (no pagination).	Not a relevant study or publication type
Nassr AA, Shazly SA, Trinidad MC, et al. Body fat index: A novel alternative to body mass index for prediction of gestational diabetes and hypertensive disorders in pregnancy. European Journal of Obstetrics, Gynecology, & Reproductive Biology 2018;228:243-248.	Not reporting a relevant outcome
Nayak AU, Vijay AMA, Indusekhar R, et al. Association of hypoglycaemia in screening oral glucose tolerance test in pregnancy with low birth weight fetus. World Journal of Diabetes 2019;10:304-310.	Not a relevant intervention
Nct. A Clinical Trial of the Effectiveness of Metformin in Addition to Usual Care in the Reduction of Gestational Diabetes Mellitus. Https://clinicaltrials.gov/show/nct02980276 2016.	Published pre-2009
Nct. Acarbose and Prandial Insulin for the Treatment of Gestational Diabetes Mellitus. Https://clinicaltrials.gov/show/nct03380546 2017.	Published pre-2009
Nct. Adding Metformin to Insulin in Controlling Pregestational and Gestational Diabetes Mellitus. Https://clinicaltrials.gov/show/nct03106870 2017.	Published pre-2009
Nct. Aerobic Exercise in Women With Gestational Diabetes. Https://clinicaltrials.gov/show/nct03067662 2017.	Published pre-2009
Nct. Carbohydrates Distribution in Pregnancy Study. Https://clinicaltrials.gov/show/nct03378908 2017.	Published pre-2009
Nct. Comparison of Insulin Alone to Insulin With Metformin to Treat Gestational Diabetes Mellitus. Https://clinicaltrials.gov/show/nct03651531 2018.	Published pre-2009
Nct. Comparison of Metformin and Insulin for Management of Gestational Diabetes Mellitus. Https://clinicaltrials.gov/show/nct03320694 2017.	Not a relevant study or publication type
Nct. Effect of Carbohydrate Distribution on Blood Glucose in Women With Gestational Diabetes Mellitus (GDM). Https://clinicaltrials.gov/show/nct03835208 2019.	Not reporting a relevant outcome
Nct. Evaluation of Exercise, in Addition to Diet, in Women With Gestational Diabetes. Https://clinicaltrials.gov/show/nct03174340 2017.	Published pre-2009

Nct. Failure Rate of GLyburide And Metformin Among Gestational Diabetics. Https://clinicaltrials.gov/show/nct03078725 2017.	Published pre-2009
Nct. Food Order Intervention for Gestational Diabetes. Https://clinicaltrials.gov/show/nct03592784 2018.	Published pre-2009
Nct. Glyburide vs Glucovance in the Treatment of GDM. Https://clinicaltrials.gov/show/nct02726490 2016.	Published pre-2009
Nct. Medical Nutrition Therapy Combined With TPF-DM in Pregnant Women With Gestational Diabetes Mellitus. Https://clinicaltrials.gov/show/nct03957603 2019.	Not reporting a relevant outcome
Nct. Metabolic Analysis for Treatment Choice in Gestational Diabetes Mellitus. Https://clinicaltrials.gov/show/nct03029702 2017.	Published pre-2009
Nct. Pregnancy Outcomes and Medical Costs According to Gestational Diabetes Mellitus Diagnostic Criteria. Https://clinicaltrials.gov/show/nct03421262 2018.	Published pre-2009
Nct. Pregnancy Outcomes: effects of Metformin Study (POEM Study). Https://clinicaltrials.gov/show/nct02947503 2016.	Published pre-2009
Nct. Screening for GDM in Women With PCOS and Controls. Https://clinicaltrials.gov/show/nct02183558 2014.	Not in a relevant population
Nct. The Acute Effect of Physical Activity on Blood Glucose in Pregnant Women. Https://clinicaltrials.gov/show/nct03644238 2018.	Not reporting a relevant outcome
Nct. The Acute Effect of Physical Activity on Postprandial Blood Glucose in Pregnant Women. Https://clinicaltrials.gov/show/nct03641170 2018.	Not reporting a relevant outcome
Nct. The Effect of Early Nutrition Intervention on the Incidence of High-risk Patients With Gestational Diabetes. Https://clinicaltrials.gov/show/nct03550976 2018.	Not reporting a relevant outcome
Nct. Threshold for Initiation of Medical Treatment for Women With Gestational Diabetes. Https://clinicaltrials.gov/show/nct03527537 2018.	Published pre-2009
Nct. Trial of Diet in Gestational Diabetes Mellitus: metabolic Consequences to Mother and Offspring. Https://clinicaltrials.gov/show/nct02244814 2014.	Published pre-2009
Nermin K, Aydin K, Duran M, et al. Effect of number of abnormal oral glucose tolerance test (OGTT) values on birthweight in women with gestational diabetes. Indian Journal of Medical Research 2013;137:95-101.	Not reporting a relevant outcome
Nevalainen J, Sairanen M, Appelblom H, et al. First-Trimester Maternal Serum Amino Acids and Acylcarnitines Are Significant Predictors of Gestational Diabetes. The Review of Diabetic Studies 2016;13:236-245.	Not a relevant intervention
Nguyen THL, Yang JW, Mahone M, et al. Are There Benefits for Gestational Diabetes Mellitus in Treating Lower Levels of Hyperglycemia Than Standard Recommendations? Canadian Journal of Diabetes 2016;40:548-554.	Not reporting a relevant outcome
Nicholson W, Bolen S, Witkop CT, et al. Benefits and risks of oral diabetes agents compared with insulin in women with gestational diabetes: A systematic review. Obstetrics and Gynecology 2009;113:193-205.	Published pre-2009
Nielsen KK, Kapur A, Damm P, et al. From screening to postpartum follow-up - the determinants and barriers for gestational diabetes mellitus (GDM) services, a systematic review. BMC Pregnancy and Childbirth 2014;14 (1) (no pagination).	Not a relevant intervention
Nilsson C, Ursing D, Strevens H, et al. Towards normalized birthweight in gestational diabetes mellitus. Acta Obstetricia et Gynecologica Scandinavica 2015;94:1090-4.	Not reporting a relevant outcome
Nishikawa T, Ono K, Hashimoto S, et al. One-hour oral glucose tolerance test plasma glucose at gestational diabetes diagnosis is a common predictor of the need for insulin therapy in pregnancy and postpartum impaired glucose tolerance. Journal of Diabetes Investigation 2018;9:1370-1377.	Not reporting a relevant outcome
Nollino L, Marcon ML, Kiwanuka E, et al. Can Nurse-Based Management Screening Ensure Adequate Outcomes in Patients With Gestational Diabetes? A Comparison of 2 Organizational Models. Quality Management in Health Care 2019;28:51-62.	Not a relevant intervention

Norgaard K, Sukumar N, Rafnsson SB, et al. Efficacy and Safety of Rapid-Acting Insulin Analogs in Special Populations with Type 1 Diabetes or Gestational Diabetes: Systematic Review and Meta-Analysis. Diabetes Therapy 2018;9:891-917.	Published pre-2009
Ntr. TANGO-DM: een studie naar afkapwaarden van zwangerschapsdiabetes. Http://www.who.int/trialsearch/trial2.aspx?Trialid=ntr7473 2018.	Not reporting a relevant outcome
Ntr. The SUGAR-DIP trial; oral medication versus insulin for diabetes in pregnancy. Http://www.who.int/trialsearch/trial2.aspx?Trialid=ntr6134 2016.	Published pre-2009
Oawada N, Aoki S, Sakamaki K, et al. Clinical significance of low result of 1-h 50-g glucose-challenge test in pregnant women. Journal of Maternal-Fetal & Neonatal Medicine 2019;32:58-61.	Published pre-2009
Odsaeter IH, Asberg A, Vanky E, et al. Hemoglobin A1c as screening for gestational diabetes mellitus in Nordic Caucasian women. Diabetology and Metabolic Syndrome 2016;8 (1) (no pagination).	Not in a relevant study type
Ogonowski J, Miazgowski T, Kuczynska M, et al. Pregravid body mass index as a predictor of gestational diabetes mellitus. Diabetic Medicine 2009;26:334-8.	Not a relevant intervention
Ogonowski J, Miazgowski T. Are short women at risk for gestational diabetes mellitus? European Journal of Endocrinology 2010;162:491-7.	Not a relevant intervention
Ogonowski J, Miazgowski T. Intergenerational transmission of macrosomia in women with gestational diabetes and normal glucose tolerance. European Journal of Obstetrics Gynecology and Reproductive Biology 2015;195:113-116.	Not reporting a relevant outcome
Ogunleye OK, Davidson KD, Gregg AR, et al. Perinatal outcomes after adopting 1- versus 2-step approach to diagnosing gestational diabetes. Journal of Maternal-Fetal and Neonatal Medicine 2017;30:186-190.	Not in a relevant population
Ojo O, Ojo OO, Wang XH, et al. The Effects of a Low GI Diet on Cardiometabolic and Inflammatory Parameters in Patients with Type 2 and Gestational Diabetes: A Systematic Review and Meta-Analysis of Randomised Controlled Trials. Nutrients 2019;11.	Not a relevant study or publication type
Okubo H, Crozier SR, Harvey NC, et al. Maternal dietary glycemic index and glycemic load in early pregnancy are associated with offspring adiposity in childhood: the Southampton Women's Survey. American Journal of Clinical Nutrition 2014;100:676-83.	Not reporting a relevant outcome
Olsen SF, Houshmand-Oeregaard A, Granstrom C, et al. Diagnosing gestational diabetes mellitus in the Danish National Birth Cohort. Acta Obstetricia et Gynecologica Scandinavica 2017;96:563-569.	Not a relevant intervention
Orecchio A, Periard D, Kashef A, et al. Incidence of gestational diabetes and birth complications in Switzerland: screening in 1042 pregnancies. Gynecological Endocrinology 2014;30:561-4.	Not in a relevant population
Osman MW, Nath M, Khalil A, et al. The effects of metformin on maternal haemodynamics in gestational diabetes mellitus: A pilot study. Diabetes Research and Clinical Practice 2018;139:170-178.	Not reporting a relevant outcome
O'Tierney-Ginn P, Presley L, Myers S, et al. Placental growth response to maternal insulin in early pregnancy. Translational Endocrinology and Metabolism 2015;100:159-165.	Not reporting a relevant outcome
Ozgu-Erdinc AS, Yilmaz S, Yeral MI, et al. Prediction of gestational diabetes mellitus in the first trimester: Comparison of C-reactive protein, fasting plasma glucose, insulin and insulin sensitivity indices. Journal of Maternal-Fetal and Neonatal Medicine 2015;28:19571962.	Not a relevant intervention
Palatnik A, Swanson K, Churchill T, et al. Association between Type of Screening for Gestational Diabetes Mellitus and Cesarean Delivery. Obstetrics and Gynecology 2017;130:539-544.	Not reporting a relevant outcome
Papastefanou I, Souka AP, Pilalis A, et al. First trimester prediction of small- and large-for-gestation neonates by an integrated model incorporating ultrasound parameters, biochemical indices and maternal characteristics. Acta Obstetricia et Gynecologica Scandinavica 2012;91:104-11.	Not a relevant intervention

Patel KR, White FV, Deutsch GH. Hepatic steatosis is prevalent in stillborns delivered to women with diabetes mellitus. Journal of Pediatric Gastroenterology and Nutrition 2015;60:152-158.	Not a relevant intervention
Patel S, Fraser A, Smith GD, et al. Associations of gestational diabetes, existing diabetes, and glycosuria with offspring obesity and cardiometabolic outcomes. Diabetes Care 2012;35:63-71.	Not a relevant intervention
Patel S, Lawlor DA, Callaway M, et al. Association of maternal diabetes/glycosuria and pre-pregnancy body mass index with offspring indicators of non-alcoholic fatty liver disease. BMC Pediatrics 2016;16 (1) (no pagination).	Not reporting a relevant outcome
Patenaude J, Lacerte G, Lacroix M, et al. Associations of Maternal Leptin with Neonatal Adiposity Differ according to Pregravid Weight. Neonatology 2017;111:344-352.	Not a relevant intervention
Paul R, Murugesh C, Chepulis L, et al. Should antenatal corticosteroids be considered in women with gestational diabetes before planned late gestation caesarean section. Australian and New Zealand Journal of Obstetrics and Gynaecology 2019;59:463-466.	Not a relevant intervention
Perez-Ferre N, Torrejon MJ, Fuentes M, et al. Association of low serum 25-hydroxyvitamin D levels in pregnancy with glucose homeostasis and obstetric and newborn outcomes. Endocrine Practice 2012;18:676-684.	Not a relevant intervention
Persson M, Winkvist A, Mogren I. Surprisingly low compliance to local guidelines for risk factor based screening for gestational diabetes mellitus - A population-based study. BMC Pregnancy & Childbirth 2009;9:53.	Not a relevant intervention
Petry CJ, Ong KK, Hughes IA, et al. Early pregnancy-associated plasma protein a concentrations are associated with third trimester insulin sensitivity. Journal of Clinical Endocrinology and Metabolism 2017;102:2000-2008.	Not a relevant intervention
Petry CJ, Ong KK, Hughes IA, et al. The influence of maternal pregnancy glucose concentrations on associations between a fetal imprinted gene allele score and offspring size at birth. BMC Research Notes 2018;11:821.	Not reporting a relevant outcome
Pintaudi B, Di Vieste G, Corrado F, et al. Improvement of selective screening strategy for gestational diabetes through a more accurate definition of high-risk groups. European Journal of Endocrinology 2014;170:87-93.	Not reporting a relevant outcome
Plasencia W, Garcia R, Pereira S, et al. Criteria for screening and diagnosis of gestational diabetes mellitus in the first trimester of pregnancy. Fetal Diagnosis & Therapy 2011;30:108-15.	Not a relevant intervention
Poolsup N, Suksomboon N, Amin M. Effect of treatment of gestational diabetes mellitus: a systematic review and meta-analysis. Plos One 2014;9:e92485.	Published pre-2009

Poolsup N, Suksomboon N, Amin M. Efficacy and safety of oral antidiabetic drugs in comparison to insulin in treating gestational diabetes mellitus: a meta-analysis. Plos One 2014;9:e109985.	Published pre-2009
Popova P, Tkachuck A, Bolotko Y, et al. Randomised controlled trial of very tight versus less tight glycaemic targets in women with gestational diabetes: preliminary results. Diabetologia 2018;61:S27-S28.	Not a relevant study or publication type
Powe CE, Allard C, Battista MC, et al. Heterogeneous contribution of insulin sensitivity and secretion defects to gestational diabetes mellitus. Diabetes Care 2016;39:1052-1055.	Not a relevant intervention
Powe CE, Nodzenski M, Talbot O, et al. Genetic determinants of glycemic traits and the risk of gestational diabetes mellitus. Diabetes 2018;67:2703-2709.	Not reporting a relevant outcome
Prieto-Sanchez MT, Ruiz-Palacios M, Blanco-Carnero JE, et al. Placental MFSD2a transporter is related to decreased DHA in cord blood of women with treated gestational diabetes. Clinical Nutrition 2017;36:513-521.	Not reporting a relevant outcome
Prutsky GJ, Domecq JP, Sundaresh V, et al. Screening for gestational diabetes: a systematic review and meta-analysis. Journal of Clinical Endocrinology and Metabolism 2013;98:4311-4318.	Published pre-2009

Prutsky GJ, Domecq JP, Wang Z, et al. Glucose targets in pregnant women with diabetes: A systematic review and meta-analysis. Journal of Clinical Endocrinology and Metabolism 2013;98:4319-4324.	Not a relevant intervention
Pugh SK, Doherty DA, Magann EF, et al. Does hypoglycemia following a glucose challenge test identify a high risk pregnancy? Reproductive Health 2009;6 (1) (no pagination).	Not a relevant intervention
Ramos GA, Hanley AA, Aguayo J, et al. Neonatal chemical hypoglycemia in newborns from pregnancies complicated by type 2 and gestational diabetes mellitus the importance of neonatal ponderal index. Journal of Maternal-Fetal and Neonatal Medicine 2012;25:267271.	Not reporting a relevant outcome
Ranasinghe PD, Maruthur NM, Nicholson WK, et al. Comparative Effectiveness of Continuous Subcutaneous Insulin Infusion Using Insulin Analogs and Multiple Daily Injections in Pregnant Women with Diabetes Mellitus: A Systematic Review and Meta-Analysis. Journal of Women's Health 2015;24:237-249.	Not in a relevant population
Refuerzo JS, Viteri OA, Hutchinson M, et al. The effects of metformin on weight loss in women with gestational diabetes: A pilot randomized, placebo-controlled trial. American Journal of Obstetrics and Gynecology 2015;212:389.e1-389.e9.	Not in a relevant population
Regnault N, Botton J, Hillier TA, et al. Higher cord C-peptide concentrations are associated with slower growth rate in the 1st year of life in girls but not in boys. Diabetes 2011;60:2152-2159.	Not reporting a relevant outcome
Renz PB, Chume FC, Timm JRT, et al. Diagnostic accuracy of glycated hemoglobin for gestational diabetes mellitus: a systematic review and meta-analysis. Clinical Chemistry & Laboratory Medicine 2019;20:20.	Not a relevant study or publication type
Retnakaran R, Ye C, Hanley A, et al. Effect of maternal gestational diabetes on the cardiovascular risk factor profile of infants at 1 year of age. Nutrition, Metabolism and Cardiovascular Diseases 2013;23:1175-1181.	Not a relevant intervention
Retnakaran R, Ye C, Hanley AJG, et al. Effect of maternal weight, adipokines, glucose intolerance and lipids on infant birth weight among women without gestational diabetes mellitus. Cmaj 2012;184:1353-1360.	Not a relevant intervention
Roeckner JT, Sanchez-Ramos L, Jijon-Knupp R, et al. Single abnormal value on 3-hour oral glucose tolerance test during pregnancy is associated with adverse maternal and neonatal outcomes: a systematic review and metaanalysis. American Journal of Obstetrics & Gynecology 2016;215:287-97.	Published pre-2009
Rowan JA, Budden A, Ivanova V, et al. Women with an HbA<inf>1c</inf> of 41-49 mmol/mol (5.9-6.6%): A higher risk subgroup that may benefit from early pregnancy intervention. Diabetic Medicine 2016;33:25-31.	Not in a relevant population
Rowan JA, Rush EC, Obolonkin V, et al. Metformin in gestational diabetes: The offspring follow-up (MiG TOFU) - Body composition at 2 years of age. Diabetes Care 2011;34:2279-2284.	Not in a relevant population
Rowe CW, Putt E, Brentnall O, et al. An intravenous insulin protocol designed for pregnancy reduces neonatal hypoglycaemia following betamethasone administration in women with gestational diabetes. Diabetic Medicine 2019;36:228-236.	Not in a relevant population
Ryan EA, Sia WW, Khurana R, et al. Glucose Control During Labour In Diabetic Women. Journal of Obstetrics and Gynaecology Canada 2012;34:1149-1157.	Not reporting a relevant outcome
Salman L, Pardo A, Krispin E, et al. Perinatal outcome in gestational diabetes according to different diagnostic criteria. Journal of Perinatal Medicine 2019;47:553-557.	Not in a relevant population
Santos LL, Santos JL, Barbosa LT, et al. Effectiveness of Insulin Analogs Compared with Human Insulins in Pregnant Women with Diabetes Mellitus: Systematic Review and Meta-analysis. Revista Brasileira de Ginecologia e Obstetricia 2019;41:104-115.	Not in a relevant population
Santos MJ, Fernandes V, Portuguese P, et al. Gestational diabetes mellitus: different management strategies should be adopted for different subsets of patients diagnosed by oral glucose tolerance test. Endocrine 2018;62:602-610.	Not in a relevant population
Sarikabadayi YU, Aydemir O, Kanmaz G, et al. Umbilical artery intima-media and wall thickness in infants of diabetic mothers. Neonatology 2012;102:157-62.	Not in a relevant population

Savona-Ventura C, Vassallo J, Craus J, et al. Biological and biochemical characteristics of a Mediterranean population with Gestational Diabetes Mellitus. Journal of Perinatal Medicine 2016;44:377-382.	Not in a relevant population
Scholtens DM, Bain JR, Reisetter AC, et al. Metabolic Networks and Metabolites Underlie Associations Between Maternal Glucose During Pregnancy and Newborn Size at Birth. Diabetes 2016;65:2039-50.	Not reporting a relevant outcome
Scholtens DM, Kuang A, Lowe LP, et al. Hyperglycemia and adverse Pregnancy Outcome follow-up study (HAPO FUS): Maternal glycemia and childhood glucose metabolism. Diabetes Care 2019;42:381-392.	Not reporting a relevant outcome
Schwartz N, Green MS, Yefet E, et al. Postprandial glycemic control during gestational diabetes pregnancy predicts the risk of recurrence. Scientific Reports 2018;8:6350.	Not in a relevant population
Senat MV, Affres H, Letourneau A, et al. Effect of Glyburide vs Subcutaneous Insulin on Perinatal Complications among Women with Gestational Diabetes: a Randomized Clinical Trial. Obstetrical & gynecological survey 2018;73:511-513.	Published pre-2009
Sesmilo G, Meler E, Perea V, et al. Maternal fasting glycemia and adverse pregnancy outcomes in a Mediterranean population. Acta Diabetologica 2017;54:293-299.	Not a relevant intervention
Seval MM, Cavkaytar S, Atak Z, et al. Should we interpret the results of 'two-step' glucose screening again according to the obstetric outcomes? Journal of Obstetrics & Gynaecology 2016;36:705-709.	Not a relevant intervention
Sevket O, Sevket A, Ozel A, et al. The use of HbA1c as an aid in the diagnosis of gestational diabetes mellitus. Journal of Obstetrics & Gynaecology 2014;34:690-2.	Not in a relevant population
Shub A, Chee T, Templeton A, et al. Timing of diagnosis of gestational diabetes and pregnancy outcomes: A retrospective cohort. Australian & New Zealand Journal of Obstetrics & Gynaecology 2019;59:96-101.	Not a relevant intervention
Simmons D, Hague WM, Teede HJ, et al. Hyperglycaemia in early pregnancy: the Treatment of Booking Gestational diabetes Mellitus (TOBOGM) study. A randomised controlled trial. Medical Journal of Australia 2018;209:405-406.	Not reporting a relevant outcome
Simmons D, Nema J, Parton C, et al. The treatment of booking gestational diabetes mellitus (TOBOGM) pilot randomised controlled trial. BMC Pregnancy and Childbirth 2018;18 (1) (no pagination).	Not reporting a relevant outcome
Simmons D, Nema J, Vizza L, et al. Treatment of booking gestational diabetes: the ToBOGM pilot randomised controlled trial. Diabetic medicine 2017;34:165-.	Published pre-2009
Singh SR, Ahmad F, Lal A, et al. Efficacy and safety of insulin analogues for the management of diabetes mellitus: a meta-analysis. CMAJ: Canadian Medical Association Journal 2009;180:385-397.	Published pre-2009
Sivalingam VN, Myers J, Nicholas S, et al. Metformin in reproductive health, pregnancy and gynaecological cancer: Established and emerging indications. Human Reproduction Update 2014;20:853-868.	Not in a relevant study type
Sklempe Kokic I, Ivanisevic M, Kokic T, et al. Acute responses to structured aerobic and resistance exercise in women with gestational diabetes mellitus. Scandinavian journal of medicine & science in sports 2018;28:1793-1800.	Not a relevant intervention
Sommer C, Sletner L, Morkrid K, et al. Effects of early pregnancy BMI, mid-gestational weight gain, glucose and lipid levels in pregnancy on offspring's birth weight and subcutaneous fat: a population-based cohort study. BMC Pregnancy & Childbirth 2015;15:84.	Not reporting a relevant outcome
Song D, Lia M, Hurley JC (2019) Recommended pre-analytical plasma glucose sampling methodology may distort gestational diabetes mellitus prevalence: implications for diagnostic thresholds. Diabetic Medicine 11:11	Not reporting a relevant outcome
Song R, Chen L, Chen Y, et al. Comparison of glyburide and insulin in the management of gestational diabetes: A meta-analysis. PLoS ONE 2017;12 (8) (no pagination).	Published pre-2009
Su DF, Wang XY. Metformin vs insulin in the management of gestational diabetes: a systematic review and meta-analysis. Diabetes research and clinical practice 2014;104:353-357.	Published pre-2009

Subramaniam A, Jauk VC, Tita A, et al. Interaction between maternal obesity and 1-hour glucose challenge test results on maternal and perinatal outcomes. American Journal of Perinatology 2015;32:771-8.	Published pre-2009
Suhaimi FA, Mohd-Yusof BN, Shariff ZM, et al. A low-gi diet improves glucose self-monitoring in women with gestational diabetes mellitus. Diabetes 2016;65:A194-.	Published pre-2009
Syed M, Javed H, Yakoob MY, et al. Effect of screening and management of diabetes during pregnancy on stillbirths. BMC public health 2011;11 Suppl 3:S2.	Not in a relevant study type
Syngelaki A, Kotecha R, Pastides A, et al. First-trimester biochemical markers of placentation in screening for gestational diabetes mellitus. Metabolism: Clinical & Experimental 2015;64:1485-9.	Not in a relevant population
Syngelaki A, Pastides A, Kotecha R, et al. First-Trimester Screening for Gestational Diabetes Mellitus Based on Maternal Characteristics and History. Fetal Diagnosis & Therapy 2015;38:14-21.	Not a relevant intervention
Tan HLE, Luu J, Caswell A, et al. Impact of new International Association of Diabetes and Pregnancy Study Groups (IADPSG) diagnostic criteria on perinatal outcomes in a regional tertiary hospital in New South Wales, Australia. Diabetes Research & Clinical Practice 2017;134:191-198.	Not in a relevant population
Tarry-Adkins JL, Aiken CE, Ozanne SE. Neonatal, infant, and childhood growth following metformin versus insulin treatment for gestational diabetes: A systematic review and meta-analysis. PLoS Medicine / Public Library of Science 2019;16:e1002848.	Published pre-2009
Telejko B, Kuzmicki M, Kretowska MZ, et al. A comparison of the International Association of Diabetes and Pregnancy Study Groups Recommendations with Former Criteria for Diagnosing Gestational Diabetes Mellitus: A Retrospective Cohort Study. Experimental and Clinical Endocrinology and Diabetes 2019;127:359-366.	Not in a relevant population
Temming LA, Tuuli MG, Stout MJ, et al. Maternal and Perinatal Outcomes in Women with Insulin Resistance. American Journal of Perinatology 2016;33:776-780.	Not a relevant intervention
Tertti K, Laine K, Ekblad U, et al. The degree of fetal metformin exposure does not influence fetal outcome in gestational diabetes mellitus. Acta diabetologica 2014;51:731-738.	Not in a relevant population
The use of dextrose/insulin infusions during labour and delivery in women with gestational diabetes mellitus: is there any point? and new zealand journal of obstetrics and gynaecology. (no pagination), 2016. Date of publication: 2016. 2016.	Not reporting a relevant outcome Australian
Tieu J, McPhee AJ, Crowther CA, et al. Screening and subsequent management for gestational diabetes for improving maternal and infant health. Cochrane database of systematic reviews (online) 2014;2014.	Published pre-2009
Tieu J, McPhee AJ, Crowther CA, et al. Screening for gestational diabetes mellitus based on different risk profiles and settings for improving maternal and infant health. Cochrane Database of Systematic Reviews 2017.	Published pre-2009
Tita ATN, Lai Y, Landon MB, et al. Predictive Characteristics of Elevated 1-Hour Glucose Challenge Test Results for Gestational Diabetes. American Journal of Perinatology 2017;34:1464-1469.	Not in a relevant population
Tonguc M, Tayyar AT, Muderris I, et al. An evaluation of two different screening criteria in gestational diabetes mellitus. Journal of Maternal-Fetal & Neonatal Medicine 2018;31:1188-1193.	Not a relevant intervention
Tyrrell J, Richmond RC, Palmer TM, et al. Genetic Evidence for Causal Relationships Between Maternal Obesity-Related Traits and Birth Weight. JAMA 2016;315:1129-40.	Not in a relevant study type
Valsamakis G, Margeli A, Vitoratos N, et al. The role of maternal gut hormones in normal pregnancy: Fasting plasma active glucagonlike peptide 1 level is a negative predictor of fetal abdomen circumference and maternal weight change. European Journal of Endocrinology 2010;162:897-903.	Not a relevant intervention
van den Berg SA, de Groot MJ, Salden LP, et al. Pregnancy diabetes: A comparison of diagnostic protocols based on point-of-care, routine and optimized laboratory conditions. Scientific reports 2015;5:16302.	Not in a relevant population

van Leeuwen M, Louwerse MD, Opmeer BC, et al. Glucose challenge test for detecting gestational diabetes mellitus: a systematic review. BJOG: An International Journal of Obstetrics & Gynaecology 2012;119:393-401.	Published pre-2009
van Leeuwen M, Opmeer BC, Yilmaz Y, et al. Accuracy of the random glucose test as screening test for gestational diabetes mellitus: a systematic review. European Journal of Obstetrics and Gynecology and Reproductive Biology 2011;154:130-135.	Published pre-2009
Vecchie A, Bonaventura A, Carbone F, et al. C-Reactive Protein Levels at the Midpregnancy Can Predict Gestational Complications. BioMed Research International 2018;2018:1070151.	Not a relevant intervention
Vesco KK, Sharma AJ, Bulkley J, et al. Association of glucose levels in pregnancy with use of health care services. Diabetes Research and Clinical Practice 2019;152:146-155.	Not a relevant intervention
Viana LV, Gross JL, Azevedo MJ. Dietary intervention in patients with gestational diabetes mellitus: a systematic review and metaanalysis of randomized clinical trials on maternal and newborn outcomes. Diabetes Care 2014;37:3345-55.	Published pre-2009
Vieira MC, McCowan LME, North RA, et al. Antenatal risk factors associated with neonatal morbidity in large-for-gestational-age infants: an international prospective cohort study. Acta Obstetricia et Gynecologica Scandinavica 2018;97:1015-1024.	Not a relevant intervention
Voormolen DN, Devries JH, Evers IM, et al. The efficacy and effectiveness of continuous glucose monitoring during pregnancy: A systematic review. Obstetrical and Gynecological Survey 2013;68:753-763.	Published pre-2009
Walsh JM, McGowan CA, Kilbane M, et al. The relationship between maternal and fetal vitamin D, insulin resistance, and fetal growth. Reproductive Sciences 2013;20:536-41.	Not in a relevant population
Walsh JM, Wallace M, Brennan L, et al. Early pregnancy maternal urinary metabolomic profile and later insulin resistance and fetal adiposity. Journal of Maternal-Fetal and Neonatal Medicine 2015;28:1697-1700.	Not in a relevant population
Wang HQ, Lai HL, Li Y, et al. The Relationship between Maternal Gestational Impaired Glucose Tolerance and Risk of LargeforGestational-Age Infant: A Meta-Analysis of 14 Studies. Journal of clinical research in pediatric endocrinology 2016;8:264-9.	Not a relevant intervention
Wang J, Chen K, Jin X, et al. Prognostic factors for cesarean section outcome of pregnant women with gestational diabetes mellitus: a systematic review and meta-analysis. Diabetes, Metabolic Syndrome and Obesity Targets and Therapy 2019;12:913-929. type	Not a relevant study or publication
Waters TP, Dyer AR, Scholtens DM, et al. Maternal and Neonatal Morbidity for Women Who Would Be Added to the Diagnosis of GDM Using IADPSG Criteria: A Secondary Analysis of the Hyperglycemia and Adverse Pregnancy Outcome Study. Diabetes Care 2016;39:2204-2210.	Not a relevant intervention
Wattar BHA, Mylrea-Lowndes B, Morgan C, et al. Use of dietary assessment tools in randomized trials evaluating diet-based pregnancy: A systematic review of literature. Current Opinion in Obstetrics and Gynecology 2016;28:455-463.	Not in a relevant study type interventions in
Wei J, Heng W, Gao J. Effects of Low Glycemic Index Diets on Gestational Diabetes Mellitus: A Meta-Analysis of Randomized Controlled Trials. Medicine 2016;95:e3792.	Published pre-2009 Clinical
Wendland EM, Torloni MR, Falavigna M, et al. Gestational diabetes and pregnancy outcomes - a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic criteria. BMC Pregnancy and Childbirth 2012;12 (no pagination).	Not a relevant intervention
Wensel TM. Role of metformin in the treatment of gestational diabetes. Annals of Pharmacotherapy 2009;43:939-943.	Published pre-2009
West J, Lawlor DA, Fairley L, et al. Differences in socioeconomic position, lifestyle and health-related pregnancy characteristics between Pakistani and White British women in the Born in Bradford prospective cohort study: The influence of the woman's, her partner's and their parents' place of birth. BMJ Open 2014;4 (6) (no pagination).	Not a relevant intervention
West J, Santorelli G, Whincup PH, et al. Association of maternal exposures with adiposity at age 4/5 years in white British and Pakistani children: findings from the Born in Bradford study. Diabetologia 2018;61:242-252.	Not reporting a relevant outcome

Whitelaw DC, Scally AJ, Tuffnell DJ, et al. Associations of circulating calcium and 25-hydroxyvitamin D with glucose metabolism in pregnancy: A cross-sectional study in European and south Asian women. Journal of Clinical Endocrinology and Metabolism 2014;99:938946.	Not a relevant intervention
Wong T, Ross GP, Jalaludin BB, et al. The clinical significance of overt diabetes in pregnancy. Diabetic Medicine 2013;30:468-474.	Not a relevant intervention
Wong VW, Lin A, Russell H. Adopting the new World Health Organization diagnostic criteria for gestational diabetes: How the prevalence changes in a high-risk region in Australia. Diabetes Research & Clinical Practice 2017;129:148-153.	Not a relevant intervention
Wroblewska-Seniuk K, Wender-Ozegowska E, Szczapa J. Long-term effects of diabetes during pregnancy on the offspring. Pediatric Diabetes 2009;10:432-40.	Not a relevant intervention
Xie GH, Zheng Z, Liu TC, et al. Health care and risk of adverse pregnancy outcomes among diabetic women: an updated metaanalysis. Archives of Gynecology and Obstetrics 2019;299:891-899.	Published pre-2009
Xu J, Ye S. Influence of low-glycemic index diet for gestational diabetes: a meta-analysis of randomized controlled trials. Journal of Maternal-Fetal & Neonatal Medicine 2018:1-6.	Not a relevant study or publication type
Xu Q, Xie Q. Long-term effects of prenatal exposure to metformin on the health of children based on follow-up studies of randomized controlled trials: a systematic review and meta-analysis. Archives of Gynecology and Obstetrics 2019;299:1295-1303.	Published pre-2009
Yamamoto JM, Benham J, Mohammad K, et al. Intrapartum glycaemic control and neonatal hypoglycaemia in pregnancies complicated by diabetes: a systematic review. Diabetic Medicine 2018;35:173-183.	Not a relevant study or publication type
Yamamoto JM, Kellett JE, Balsells M, et al. Gestational diabetes mellitus and diet: A systematic review and meta-analysis of randomized controlled trials examining the impact of modified dietary interventions on maternal glucose control and neonatal birth weight. Diabetes Care 2018;41:1346-1361.	Not a relevant study or publication type
Yamashita H, Yasuhi I, Fukuda M, et al. The association between maternal insulin resistance in mid-pregnancy and neonatal birthweight in uncomplicated pregnancies. Endocrine Journal 2014;61:1019-1024.	Not a relevant intervention
Yamashita H, Yasuhi I, Kugishima Y, et al. Factors associated with patients with gestational diabetes in Japan being at increased risk of requiring intensive care. International Journal of Gynecology and Obstetrics 2018;140:170-174.	Not a relevant intervention
Yeral MI, Ozgu-Erdinc AS, Uygur D, et al. Prediction of gestational diabetes mellitus in the first trimester, comparison of fasting plasma glucose, two-step and one-step methods: a prospective randomized controlled trial. Endocrine 2014;46:512-518.	Not in a relevant study type
Yilmaz H, Cakmak M, Darcin T, et al. Elevated plasma levels of betatrophin in women with gestational diabetes mellitus. Experimental and Clinical Endocrinology and Diabetes 2015;123:376-381.	Published pre-2009
Zawiejska A, Wender-Ozegowska E, Grewling-Szmit K, et al. Short-term antidiabetic treatment with insulin or metformin has a similar impact on the components of metabolic syndrome in women with gestational diabetes mellitus requiring antidiabetic agents: Results of a prospective, randomised study. Journal of Physiology and Pharmacology 2016;67:227-233.	Not in a relevant population
Zawiejska A, Wender-Ozegowska E, Radzicka S, et al. Maternal hyperglycemia according to IADPSG criteria as a predictor of perinatal complications in women with gestational diabetes: A retrospective observational study. Journal of Maternal-Fetal and Neonatal Medicine 2014;27:1526-1530.	Not reporting a relevant outcome
Zeng YC, Li MJ, Chen Y, et al. The use of glyburide in the management of gestational diabetes mellitus: a meta-analysis. Advances in Medical Sciences 2014;59:95-101.	Published pre-2009
Zhang R, Han S, Chen GC, et al. Effects of low-glycemic-index diets in pregnancy on maternal and newborn outcomes in pregnant women: a meta-analysis of randomized controlled trials. European Journal of Nutrition 2018;57:167-177.	Published pre-2009
Zhao LP, Sheng XY, Zhou S, et al. Metformin versus insulin for gestational diabetes mellitus: A meta-analysis. British Journal of Clinical Pharmacology 2015;80:1224-1234.	Published pre-2009

Zhen XM, Li X, Chen C. Longer-term outcomes in offspring of GDM mothers treated with metformin versus insulin. Diabetes Research and Clinical Practice 2018;144:82-92.	Published pre-2009
Zhen XM, Li X, Chen C. Metformin versus insulin for gestational diabetes: The reporting of ethnicity and a meta-analysis combining English and Chinese literatures. Obesity Medicine 2018;11:48-58.	Published pre-2009
Zhu B, Zhang L, Fan YY, et al. Metformin versus insulin in gestational diabetes mellitus: a meta-analysis of randomized clinical trials. Irish Journal of Medical Science 2016;185:371-381.	Published pre-2009
Zhu Y, Olsen SF, Mendola P, et al. Growth and obesity through the first 7 y of life in association with levels of maternal glycemia during pregnancy: a prospective cohort study. American Journal of Clinical Nutrition 2016;103:794-800.	Not in a relevant population
Zilberlicht A, Feferkorn I, Younes G, et al. The mutual effect of pregestational body mass index, maternal hyperglycemia and gestational weight gain on adverse pregnancy outcomes. Gynecological Endocrinology 2016;32:416-20.	Not in a relevant population

Appendix 3 — Summary of individual studies

Data Extraction

Question 1: What are the risks of short and long-term adverse outcomes in the newborn associated with incremental increases in maternal glucose level?

Table 46: Born in Bradford IPD (Farrar 2016 Chapter 2)

Study Reference	**Born in Bradford IPD (Farrar 2016 Chapter 2)**
Study Design	Design A prospective birth cohort study Objective To establish the nature of the association of fasting and post-load glucose levels with adverse perinatal outcomes in a large cohort of SA women and compare those findings with a similarly sized cohort of WB women Dates NR Country UK Setting Bradford Royal Infirmary

Patient recruitment and eligibility

Recruitment: At OGTT appointment offered to all women booked for delivery at Bradford Royal Infirmary

Inclusion criteria: Women who delivered a live singleton baby at the Bradford Royal Infirmary, Bradford, UK

Exclusion criteria: Women who did not deliver at Bradford Royal Infirmary, had a multiple pregnancy, stillbirth, existing diabetes were excluded from the study. In addition, women who had missing data (baseline questionnaire, OGTT or ethnicity), were diagnosed with GDM were excluded from the analysis **Population Characteristics** **Other:**

Sample size

N screened/invited = 13,773
N eligible = 13,061
N enrolled = NR
N excluded (with reason) = 2243 (did not complete baseline questionnaire), 444 (did not complete OGTT) and 21 (missing data on ethnicity) N lost to follow-up = NR
N completed = 10,353
N excluded from analysis = 844 (diagnosed with GDM)

N included in analysis = 9509 (WB=3888, SA=4821, other=800) Maternal demographics

Characteristic	All women: mean (SD) or n (%)	Number of patients with available data
Maternal age at delivery, years	27.3 (5.5)	9509
Aged ≥35 years	1092 (11.5)	–
Cardiometabolic health		

Pre-pregnant BMI, kg/m^2	*NR*	–
BMI at booking, kg/m^2	25.8 (5.6)	9073
Obese (BMI ≥30 kg/m^2)	1808 (19.9)	–
Weight, kg	*NR*	–
Ethnicity, n (%)		
White British	3888 (40.9)	9509
Black	*NR*	–
South Asian	4821 (50.7)	9509
East Asian	*NR*	–
Mixed	*NR*	–
Other	800 (8.4)	9509
Medical history/risk factors, n (%)		
Family history of hypertension	2519 (27.4)	9203
Family history of diabetes	2313 (25.1)	9212
Smoking status		9494
Never	6518 (68.7)	–
Pre-pregnancy	1359 (14.3)	–
In pregnancy	1617 (17.0)	–
Any alcohol during pregnancy	1950 (20.6)	9477
Obstetric history, n (%)		
Nulliparous	3813 (41.7)	9151
Previous GDM[a]	56 (1.1)	5338
Previous macrosomia (≥4 kg)[a]	359 (8.0)	4464
		9383
Education level, n (%) *<5 GCSEs* *≥5 GCSEs*	2024 (21.6)	–
	2954 (31.5)	–
A level	1389 (14.8)	–
Higher than A level	2402 (25.6)	–
Other	614 (6.5)	–

[a]Percentages relate to multiparous women only (N=5345)
For maternal age and maternal BMI, the values are mean (SD); for all other variables (that are categorical) the values are n (%) Maternal glycaemic characteristics

Glucose tolerance	All women: mean (SD), median (IQR) or n (%)	Number of patients with available data
FPG, mmol/L	4.4 (4.2–4.7)	9509
75g OGTT, mmol/L	*NR*	*NR*
1 hour	*NR*	*NR*
2 hours	5.4 (4.7–6.1)	9509
3 hours	*NR*	*NR*
Gestational age at OGTT (weeks)	26.3 (1.9)	9509

For maternal gestational age at OGTT the values are mean (SD); for maternal gestational fasting and post-load glucose levels, values are median (IQR)

Baseline characteristics were also reported separately by ethnicity (WB, SA and other)

Methods

Duration of follow-up
NR

Method of blood glucose measurement
75g OGTT, comprising fasting and 2-hour post-load samples after an overnight fast, offered at around 26–28 weeks' gestation

Diagnostic criteria and test(s)
GDM was defined according to modified WHO criteria operating at the time: either FPG ≥6.1 mmol/L or 2 h 75g OGTT ≥7.8 mmol/L

Glucose category cut-offs

- FPG level – category 1, <4.3 mmol/L; category 2, 4.3–4.4 mmol/L; category 3, 4.5–4.7 mmol/L; category 4, 4.8–4.9 mmol/L; category 5, 5.0–5.2 mmol/L; category 6, 5.3–5.6 mmol/L; category 7, 5.7–6.0 mmol/L
- Post-load plasma glucose level – category 1, <4.7 mmol/L; category 2, 4.7–5.4 mmol/L; category 3, 5.5–6.2 mmol/L; category 4, 6.3–6.6 mmol/L; category 5, 6.7–7.2 mmol/L; category 6, 7.3–7.5 mmol/L; category 7, 7.6–7.7 mmol/L

Outcomes

Primary endpoint

- LGA, defined as BW of >90th percentile for gestational age when BW was converted into SD scores standardised for gestational age and gender relative to the UK-WHO growth standard
- Infant adiposity, defined as sum of skinfolds >90th percentile for gestational age. Skinfold thickness (triceps and subscapular) were summed and the 90th percentile was established from quantile regression using six gender–ethnic groups (combining gender and ethnic origin) and adjusted for parity (0, 1, 2, 3+)
- C-section, abstracted from medical records

Secondary endpoints

- Preeclampsia, abstracted from medical records and defined as new-onset proteinuria (>300 g in 24 hours) together with blood pressure of ≥140/90 mmHg after 20 weeks' gestation on more than one occasion
- Preterm delivery, abstracted from medical records
- Shoulder dystocia, abstracted from medical records
- Instrumental vaginal delivery, abstracted from medical records
- Admission to the neonatal unit, abstracted from medical records

Pregnancy outcomes

Outcome	All women: mean (SD), or n (%)	Number of patients with available data
Gestational age at delivery (weeks)	39.7 (1.7)	9509
Male gender	4884 (51.4)	9509
Pre-term birth (<37 weeks)	471 (5.0)	9509

Pregnancy complications		
Preeclampsia	229 (2.5)	9120
Stillbirth	*NR*	*NR*

For gestational age at delivery, the values are mean (SD); for all other variables (that are categorical) the values are n (%)

Adverse neonatal outcomes

Neonatal outcomes according to maternal glycaemic status

Outcome	All women: n (%)	Number of patients with available data
Perinatal mortality	*NR*	*NR*
Mode of birth	*NR*	*NR*
Induction of labour	*NR*	*NR*
Vaginal delivery	*NR*	*NR*
Instrumental delivery	(12.4)	7519
C-section (unspecified if emergency or planned)	(20.9)	9509
Macrosomia	*NR*	*NR*
LGA	*NR*	*NR*
BW of >90th percentile	(6.2)	9508
Sum of skinfolds of >90th percentile	(10.6)	6458
Birth injury	*NR*	*NR*
Shoulder dystocia	(1.4)	7526
Brachial plexus neuropathy	*NR*	*NR*
Neonatal hypoglycaemia	*NR*	*NR*
Admission to NICU	(4.3)	9509
Long-term outcomes	*NR*	*NR*
Greater adiposity	*NR*	*NR*
Cardiometabolic ill-health	*NR*	*NR*

OR of primary outcomes by FPG category relative to baseline category

	All women (N=9509)	
Outcome	**OR**	**95% CI**
BW of >90th percentile (i.e. LGA)		
1 (<4.3 mmol/L, reference)	1.00	–
2 (4.3–4.4 mmol/L)	1.18	0.90–1.54

3 (4.5–4.7 mmol/L)	1.35	1.04–1.74
4 (4.8–4.9 mmol/L)	1.42	1.02–1.97
5 (5.0–5.2 mmol/L)	1.90	1.35–2.67
6 (5.3–5.6 mmol/L)	3.10	2.00–4.79
7 (5.7–6.0 mmol/L)	2.60	1.35–5.04
Sum of skinfolds of >90th percentile		
1 (<4.3 mmol/L, reference)	1.00	–
2 (4.3–4.4 mmol/L)	1.11	0.88–1.40
3 (4.5–4.7 mmol/L)	1.40	1.14–1.72
4 (4.8–4.9 mmol/L)	1.61	1.24–2.09
5 (5.0–5.2 mmol/L)	2.02	1.54–2.64
6 (5.3–5.6 mmol/L)	3.23	2.29–4.56
7 (5.7–6.0 mmol/L)	2.73	1.53–4.87
C-section		
1 (<4.3 mmol/L, reference)	1.00	–
2 (4.3–4.4 mmol/L)	0.98	0.84–1.13
3 (4.5–4.7 mmol/L)	1.11	0.96–1.28
4 (4.8–4.9 mmol/L)	1.17	0.97–1.41
5 (5.0–5.2 mmol/L)	1.20	0.98–1.48
6 (5.3–5.6 mmol/L)	1.14	0.84–1.55
7 (5.7–6.0 mmol/L)	2.14	1.34–3.41

OR of primary outcomes by 2 h post-load 75g OGTT category relative to baseline category

Outcome	All women (N=9509)	
	OR	95% CI
BW of >90th percentile		
1 (<4.7 mmol/L, reference)	1.00	–
2 (4.7–5.4 mmol/L)	0.95	0.74–1.23
3 (5.5–6.2 mmol/L)	1.08	0.83–1.39
4 (6.3–6.6 mmol/L)	1.29	0.92–1.80
5 (6.7–7.2 mmol/L)	1.58	1.14–2.19
6 (7.3–7.5 mmol/L)	1.71	1.04–2.81
7 (7.6–7.7 mmol/L)	1.29	0.65–2.60

Sum of skinfolds of >90th percentile		
1 (<4.7 mmol/L, reference)	1.00	–
2 (4.7–5.4 mmol/L)	1.02	0.81–1.29
3 (5.5–6.2 mmol/L)	1.32	1.05–1.65
4 (6.3–6.6 mmol/L)	1.84	1.40–2.41
5 (6.7–7.2 mmol/L)	1.94	1.47–2.55
6 (7.3–7.5 mmol/L)	2.29	1.54–3.39
7 (7.6–7.7 mmol/L)	2.53	1.53–4.17
C-section		
1 (<4.7 mmol/L, reference)	1.00	–
2 (4.7–5.4 mmol/L)	0.95	0.82–1.10
3 (5.5–6.2 mmol/L)	1.07	0.92–1.24
4 (6.3–6.6 mmol/L)	1.11	0.91–1.36
5 (6.7–7.2 mmol/L)	1.00	0.81–1.23
6 (7.3–7.5 mmol/L)	1.31	0.96–1.79
7 (7.6–7.7 mmol/L)	1.15	0.76–1.74

OR per mmol/L of glucose or per SD of glucose on FPG

Outcome	All women (N=9509)	
	OR	95% CI
BW of >90th percentile	1.31	1.20–1.43
Sum of skinfolds of >90th percentile	1.35	1.25–1.45
C-section	1.09	1.03–1.15

OR per mmol/L of glucose or per SD of glucose on 2 h post-load 75g OGTT

Outcome	All women (N=9509)	
	OR	95% CI
BW of >90th percentile	1.17	1.07–1.29
Sum of skinfolds of >90th percentile	1.31	1.21–1.42
C-section	1.05	0.99–1.11

Thresholds of fasting and post-load glucose levels (mmol/L) that would identify an OR of ≈1.75 for BW of >90th percentile and sum of skinfolds of >90th percentile

Outcome	All women (N=10,356)	
	FPG	2 h 75g OGTT
BW of >90th percentile	5.3	Not possible to determine
Sum of skinfolds of >90th percentile	5.2	7.5
Average glucose level for both BW and sum of skinfolds of >90th percentile	5.3	7.5

Results were also reported separately by ethnicity (WB, SA and other). Frequency of primary outcomes across glucose categories by ethnicity is reported graphically

Authors' Conclusions	Results of the study are compared with the IADPSG analysis of the HAPO study, mainly focusing on GDM diagnostic criteria, the prevalence of GDM and the effect of ethnicity on the results

Abbreviations: BMI, body mass index; BW, body weight; CI, confidence interval; FPG, fasting plasma glucose; GCSEs, general certificate of secondary education; GDM, gestational diabetes mellitus; LGA, large for gestational age; NICU, neonatal intensive care unit; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; SA, South Asian; WB, white British, whole blood; WHO, World Health Organization

Table 47: Farrar 2016 (1678) Chapter 3 SLR

Study Reference	Farrar 2016 (1678) Chapter 3 SLR
Study Design	Design Systematic literature review Objective To determine associations between fasting and post-load glucose levels, and both perinatal and longer-term maternal and offspring outcomes. Search dates Any date until March 2013, updated on 16th September 2013 and 20th October 2014. Country Various Setting NR
	Study eligibility **Inclusion (PICOS)**

UK NSC external review

Population Characteristics

Population – Screening for G	Pregnant women who had undergone assessment of glucose tolerance using an OGTT, including the 75 g and 100 g stational Diabetes tests, or 50 g OGCT
Intervention	N/A
Comparator	N/A
Outcomes	Outcome data reported as numbers of events in each of two or more defined glucose categories, as ORs or risk ratios in each category relative to a specified baseline category, or as ORs or risk ratios per SD or per 1 mmol/L of glucose. Studies had to report at least one of the following outcomes: • Perinatal maternal outcomes ○ C-section (elective or emergency) ○ Induction of labour ○ Instrumental (assisted delivery) (ventouse or forceps) ○ Pregnancy-induced hypertension (however defined) ○ Pre-eclampsia (however defined) • Perinatal infant outcomes ○ Macrosomia (BW of ≥4.0 kg).

Study Reference	Farrar 2016 (1678) Chapter 3 SLR	
		○ LGA (BW of ≥90th percentile, or however defined) ○ Preterm birth (<37 weeks' gestation) ○ Birth injury/trauma: ✦ Shoulder dystocia ✦ Erb's palsy ✦ Fractured clavicle ○ Admission to special care or higher-care facility ○ Neonatal hypoglycaemia • Longer-term maternal or offspring outcomes ○ Type 2 diabetes (offspring or mother) ○ Cardiovascular disease (offspring or mother) ○ Obesity (offspring or mother) (however defined)
	Study design	Published and ongoing cohort studies and control (placebo or no active treatment) arms of randomised trials

Exclusion

- Women with pre-existing diabetes or treated GDM
- Studies of intravenous glucose testing

Other

NR

Flow of Studies (PRISMA)

Characteristic	Details
Design	Not summarised
Sample sizes	Not summarised

Study Reference **Farrar 2016 (1678) Chapter 3 SLR**

Setting and timing	Not summarised

- Database results: 15,916
- Records after duplicates removed: 11,219
- Hand-searches/other sources: 22
- Title/abstracts reviewed: 11,241
- Full-texts reviewed: 125
- Cohorts with IPD: 2
- Articles included in qualitative synthesis: 57
- Articles included in quantitative synthesis (meta-analysis): 37

Included study characteristics

Participants	Not summarised
Diagnostic criteria for GDM	Variety of criteria used, including Carpenter and Coustan or NDDG, WHO or defined in study
Interventions and comparisons	N/A
Outcomes	Studies eporting on: associations between glucose levels (from OGTT or OGCT) r split into three or more categories and adverse perinatal outcomes (28 studies) • associations between glucose levels (from OGTT or OGCT) split into two categories with adverse perinatal outcomes (20 studies) • longer-term outcomes in either mother or offspring (5 studies) • no numerical data (5 studies) • results from a 75 g OGTT in a non-fasted population (1 study)
Funding	NR
Conflicts of interest	NR

Definition of GDM

As defined in the individual study

Study Reference	Farrar 2016 (1678) Chapter 3 SLR
Methods	Searches Sources searched • MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid SP) • EMBASE (via Ovid SP) • Cumulative Index to Nursing and Allied Health Literature (CINAHL) Plus (via EBSCO*host*) • CENTRAL (via The Cochrane Library/Wiley Interscience) • Cochrane Database of Systematic Reviews (CDSR) (via The Cochrane Library/Wiley Interscience) • Database of Abstracts of Reviews of Effects (DARE) (via The Cochrane Library/Wiley Interscience) • Health Technology Assessment (HTA) database (via The Cochrane Library/Wiley Interscience) • NHS Economic Evaluation Database (NHS EED) (via The Cochrane Library/Wiley Interscience) • Cochrane Methodology Register (CMR) (via The Cochrane Library/Wiley Interscience) • Reference searches of included journal articles and related systematic reviews Screening and selection process All records (title, publication details and abstracts if available) were screened for eligibility, independently, by two reviewers. Records previously identified by the March 2013 search were rescreened again to ensure that the screening standard was high and consistent across all searches. All studies identified as potential 'includes' were checked by a second reviewer. Disagreements were resolved by discussion or by a third reviewer. Quality assessment Risk of bias was assessed using a modified version of the Critical Appraisal Skills Programme (CASP) and Quality in Prognosis Studies (QUIPS) quality assessment tool. The following quality criteria were considered: • representative nature of included population • loss to follow-up • consistency of glucose measurement and outcome assessment

Study Reference **Farrar 2016 (1678) Chapter 3 SLR**

- blinding of participants and medical practitioners to glucose level
- blinding of outcome assessors to glucose level
- selective reporting of outcomes
- adjustment of results for key confounding variables

Each criterion was classed as being at low, high or unclear risk of bias. One reviewer performed the quality assessment; all assessments were then checked by a second reviewer.

Contact with authors and individual participant data

Two eligible cohorts with IPD were included: the BiB study and the Atlantic Diabetes in Pregnancy cohort (Atlantic DIP). When outcomes were not reported explicitly in the data set they were derived from available data if possible (e.g., macrosomia, LGA and preterm birth were calculated from BW and gestational age data).

Statistical analyses

General approach: Statistical analyses were based on the number of women, and number of outcome events in each glucose category in each study. For the BiB and Atlantic DIP cohorts, glucose levels were divided into seven categories, with equal numbers of women in each category; for other published eligible studies, the categories set in the study were used. Studies that did not report outcomes by glucose categories were not included in these unadjusted analyses of outcome risk by glucose category. Within each glucose category the risk was calculated by dividing the number of outcome events by the total number of women in that category. Before modelling the identified associations and pooling results from studies, risk per glucose category was graphed where possible against the categories to assess the shape of the association for linearity. In studies that reported adjusted ORs or risk ratios for each glucose category, these results were similarly plotted to check the shape of the association and identify any divergence from results using unadjusted data.

Additional details on the statistical approach used for studies reporting odds ratios or risk ratios per SD or 1 mmol/L of glucose, studies reporting three or more glucose categories and cohorts with individual participant data are provided in the full-text.

OR per 1 mmol/L increases of glucose – pregnancy outcomes

Study Reference Outcome, OR (95% CI) Farrar 20	16g OGCT (1678) Chapter		SLR g OGTT			g OGTT			Combined OGTT (75 g and 100 g)		
	1h	2h	Fasting	1h	2h	Fasting	1h	2h	Fasting	1h	2h
Gestational age at birth	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Pre-term birth	1.06 (0.96–1.17)	NR	0.77 (0.62–0.96)	NR	1.07 (1.00–1.15)	NR	NR	0.87 (0.41–1.87)*	0.77 (0.62–0.96)	NR	1.07 (0.99–1.15)
Pregnancy complications	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Pre-eclampsia	1.25 (1.13–1.39)	NR	2.37 (1.40–4.04)	1.19 (1.15–1.23)*	1.22 (1.14–1.30)	1.40 (0.85–2.31)*	NR	1.37 (1.14–1.65)*	2.15 (1.45–3.19)	1.19 (1.15–1.24)	1.23 (1.18–1.29)
PIH/pre-eclampsia	1.02 (0.75–1.38)*	NR	2.00 (1.23–3.23)	NR	1.21 (1.08–1.35)	1.29 (0.77–2.16)*	NR	1.14 (0.96–1.35)*	1.91 (1.49–2.43)	NR	1.19 (1.08–1.30)
Stillbirth	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Adverse Neonatal Outcomes

*Value based on one study only.

OR per 1 mmol/L increases of glucose – neonatal outcomes

Outcome, OR (95% CI)	50 g OGCT	75 g OGTT			100 g OGTT		Combined OGTT (75 g and 100 g)		
	1h	Fasting	1h	2h	Fasting	2h	Fasting	1h	2h
Perinatal mortality	NR	NR	NR	NR	NR	NR	NR	NR	NR
Mode of birth	NR	NR	NR	NR	NR	NR	NR	NR	NR
Induction of labour	NR	1.39 (1.06–1.82)	NR	1.11 (1.03–1.19)	NR	NR	1.31 (1.14–1.50)	NR	1.10 (1.04–1.16)
Vaginal delivery	NR	NR	NR	NR	NR	NR	NR	NR	NR
Instrumental delivery	1.14 (1.04–1.24)	0.99 (0.78–1.25)	NR	1.09 (1.02–1.17)	NR	NR	0.99 (0.87–1.13)	NR	1.07 (1.03–1.12)
C-section if (unspecified	1.35 (1.23–1.49)	1.66 (1.52–	1.18	1.10 (0.98–	1.25 (1.03–1.51)*	0.95 (0.72–1.25)*	1.59 (1.49–	1.18 (1.15–	1.10 (0.96–1.25)

Farrar 2016 (1678) Chapter 3 SLR

emergency or planned)		1.82)	(1.15–1.21)*	1.24)	1.25 (0.96–1.64)*	1.17 (1.06–1.29)*	1.70)	1.20)	
Macrosomia	1.14 (1.10–1.18)	1.96 (1.57–2.43)	NR	1.19 (1.14–1.25)	1.99 (1.62–2.44)*	1.26 (1.12–1.41)*	2.06 (1.86–2.28)	NR	1.21 (1.16–1.26)
					2.69 (1.94–3.72)*	1.63 (1.16–2.31)*			
LGA	1.32 (1.19–1.46)	2.15 (1.60–2.91)	1.24 (1.20–1.27)*	1.20 (1.13–1.28)	1.89 (1.11–3.21)*	1.33 (1.13–1.55)*	2.11 (1.73–2.58)	1.24 (1.20–1.27)	1.22 (1.19–1.25)
						1.44 (1.04–2.00)*			
Birth injury	NR	NR	NR	NR	NR	NR	NR	NR	NR
Shoulder dystocia	1.26 (1.10–1.43)	1.92 (1.29–2.85)	NR	1.41 (1.03–1.92)	2.38 (0.81–7.01)*	1.61 (1.25–2.08)*	1.97 (1.36–2.85)	NR	1.38 (1.22–1.56)
						0.81 (0.26–2.54)*			
Brachial plexus neuropathy	NR	NR	NR	NR	NR	NR	NR	NR	NR
Neonatal hypoglycaemia	1.38 (1.00–1.92)	1.37 (1.20–1.57)	1.07 (1.03–1.10)*	1.13 (1.09–1.18)	NR	1.09 (0.66–1.80)*	1.37 (1.20–1.57)	NR	1.13 (1.09–1.18)
Admission to NICU	NR	NR	NR	NR	NR	NR	NR	NR	NR
Long-term outcomes	NR	NR	NR	NR	NR	NR	NR	NR	NR
Greater adiposity	NR	NR	NR	NR	NR	NR	NR	NR	NR
Cardiometabolic ill-health	NR	NR	NR	NR	NR	NR	NR	NR	NR

*Value based on one study only. No meta-analysis was carried out due to the limited number of studies.

OR per glucose tolerance test result – pregnancy outcomes

Outcome, OR (95% CI)	Negative OGCT versus positive OGCT	No elevated OGTT versus one elevated OGTT result

Study Reference **Farrar 2016 (1678) Chapter 3 SLR**

Gestational age at birth	NR	NR
Pre-term birth	1.08 (0.80–1.47)	0.91 (0.42–1.97)*
		1.00 (0.50–2.00)*
		1.44 (0.43–4.80)*

Study Reference

Pregnancy complications	NR	NR
Pre-eclampsia	1.26 (1.10–1.44)	0.57 (0.14–2.25)*
Stillbirth	NR	NR

*Value based on one study only. No meta-analysis was carried out due to the limited number of studies.

OR per glucose tolerance test result – neonatal outcomes

Outcome, OR (95% CI)	Negative OGCT versus positive OGCT	No elevated OGTT versus one elevated OGTT result
Perinatal mortality	NR	NR
Mode of birth	NR	NR
Induction of labour	0.80 (0.61–1.05)	NR
Vaginal delivery	NR	NR
Instrumental delivery	NR	0.96 (0.46–2.00)*
C-section (unspecified if emergency or planned)	1.27 (1.21–1.34)	1.74 (1.12–2.71)*
		2.30 (1.67–3.17)*
		0.68 (0.34–1.32)*
		1.48 (0.99–2.21)*
		1.72 (0.88–3.37)*
Macrosomia	1.34 (1.13–1.59)	1.13 (0.59–2.19)
		2.83 (1.18–6.78)
		2.16 (1.00–4.69)
LGA	1.42 (1.24–1.63)	1.99 (1.07–3.71)*
Birth injury	NR	NR
Shoulder dystocia	2.79 (1.30–6.01)*	0.21 (0.02–1.82)*
Brachial plexus neuropathy	NR	NR
Neonatal hypoglycaemia	1.44 (0.34–6.07)	1.32 (0.50–3.45)*
		1.17 (0.20–6.94)*
		1.41 (0.77–2.60)*

Farrar 2016 (1678) Chapter 3 SLR

Admission to NICU	NR	NR
Long-term outcomes	NR	NR
Greater adiposity	NR	NR
Cardiometabolic ill-health	NR	NR

*Value based on one study only. No meta-analysis was carried out due to the limited number of studies.

Quality Assessment

Study	Representative population	Loss to follow-up	Measurement of consistent glucose	Measurement of consistent outcome	'Blinding' of glucose measurements	'Blinding' of outcomes	Selective reporting	Adjusted results presented
Aberg 2001	Low	Low	Low	Low	Unclear	Unclear	Low	High
Aris 2014	Low	Low	Low	Low	Unclear	Unclear	Low	Low
Atlantic DIP	Low	Low	Low	Low	Low	Low	Low	Low

Study Reference	Farrar 2016 (1678) Chapter 3 SLR								
	BiB	Low	Low	Low	Low	Low	Low	Low	Low
	Black 2010	Low	Low	Low	Low	High	High	Low	Low
	Carr 2011	Low/ moderate	Low	Low	Low	High	High	Low	Low
	Chadna 2006	Unclear	Low	Unclear	Unclear	High	High	Unclear	High
	Cheng 2007	Low	Low	Low	Unclear	High	High	Unclear	Low
	Dudhbhai 2006	Low	Low	Low	Low	High	High	Low	High
	Figueroa 2013	Low (but subset of trial)	Low	Low	Low	Unclear	Unclear	Low	Low
	Forest 1994	Low	Low	Low	Low	High	High	Low	High
	Franks 2006	High (Pima Indian)	High	Low	Unclear	Unclear	Unclear	Low	Limited adjustment

Study Reference	Farrar 2016 (1678) Chapter 3 SLR								
	HAPO 2009	Low	Low	Low	Low	Low	Low	Low	Low
	HAPO 2008	Low	Low	Low	Low	Low	Low	Low	Low

Study Reference								
HAPO 2010	Low	Low	Low	Low	Low	Low	Low	Low
Hedderson 2003	Low	Low	Low	Low	High	High	Low	High
Herman 1988	Low	Low	Low	Unclear	High	High	High	High
Hillier 2007	Low	Low	Low	Low	High	High	Low	Unclear
Hillier 2008	Low	Low	Low	Unclear	Unclear	Unclear	Unclear	Low
Jensen 2001	High (higher-risk group)	Low	Low	Low	High	High	Low	High
Jensen 2008	High (higher-risk group)	Low	Low	Low	High	High	Low	High
Jiménez-Moleón 2002	Low	Low	Low	Low	High	High	Low	High
Kerényi 2009	Unclear	Low	Low	Low	Low	Unclear	Unclear	Unclear
Khan 1994	Unclear/ high-risk (Pakistani population)	Low	Low	Low	High	High	Unclear	High

Khoshniat 2010	Unclear (Iranian)	Low	Low	Unclear	Unclear	Unclear	Unclear	High
Landon 2011	Low (but subset of trial)	Low	Low	Low	Unclear	Low	Low	Low
Langer 2005	Low	Unclear	Low	Low	High	High	Low	High
Lapolla 2007	Low	High	Low	Low	Unclear	Unclear	Low	High
Lao 2003	Low (Chinese)	Low	Low	Low	High	High	Low	High
Little 1990	Low	Low	Low	Unclear	Unclear	Unclear	Unclear	High
Lurie 1998	Low	Low	Low	Low	Unclear	Low	Low	High
Ma 2013	Low	Unclear	Low	Low	High	Low	Unclear	High
Metzger 2010	Low	Low	Low	Low	Low	Low	Low	Low
Moses 1995	Low	Unclear	Low	Low	Unclear	Unclear	Low	High
Naylor 1996	Low	Low	Low	Unclear	Low	Unclear	Unclear	High

Study Reference
Farrar 2016 (1678) Chapter 3 SLR

Nord 1995	Unclear	Low	Low	Unclear	Unclear	Unclear	Unclear	High
Ong 2008	Low	Low	Low	Unclear	High	High	Unclear	High
Özekinci 2011	Low	Low	Low	Low	High	High	Unclear	High
Pettitt 1980	High (Pima Indian)	Low	Low	Low	Unclear	Unclear	Unclear	High
Pettitt 1991	High (Pima Indian)	Unclear	Low	Low	Unclear	Unclear	Low	Limited adjustment
Pettitt 2010	Low	Low	Low	Low	Unclear	Unclear	Low	Low
Pugh 2010	Low	Low	Low	Unclear	Unclear	Unclear	Unclear	Low
Retnakaran 2008	Low	Low	Low	Low	Unclear	Unclear	Low	High
Riskin-Mashiah 2009	Low	Low	Low	Low	High	High	Low	Limited adjustment
Savona-Ventura 2010	Low	Low	Low	Unclear	High	High	Unclear	High
Scholl 2001	Low	Low	Low	Low	Unclear	Unclear	Low	Low
Sermer 1995	Low	Low	Low	Low	Low	Low	Low	High

Stamilo 2004	Low	Low	Low	Low	High	High	Low	Low
Subramaniam 2014	Low	Low	Low	Unclear	High	High	Unclear	Low
Tallarigo 1986	Low	Low	Low	Low	Unclear	Unclear	Low	High
Tarim 2011	Low	Low	Low	Low	High	High	Unclear	High
Vambergue 2000	Low	Low	Low	Low	Unclear	Unclear	Low	High
Wang 2013	Low	Low	Low	Low	High	High	Low	Low
Witter 1988	Low, but young age group	Low	Low	Low	High	High	Low	High
Yee 2011	Low	Low	Low	Low	High	High	Low	Low
Yogev 2005	Low	Low	Low	Low	Unclear	Unclear	Low	High

Authors' Conclusions	Across the whole spectrum of glucose levels there was an increasing risk for the majority of reported adverse perinatal outcomes including macrosomia, LGA, C-section, pre-eclampsia, neonatal hypoglycaemia and shoulder dystocia. Associations between risk of an outcome and graded increases in glucose level seemingly applied to all glucose loads (50 g, 75 g and 100 g) and at all measurement times (fasting, and 1-hour and 2-hour post load), although the strength of these associations varied. Associations were stronger for fasting glucose levels than post-load glucose levels and for the 75 g OGTT compared with the 100 g OGTT.
Study Reference	**Farrar 2016 (1678) Chapter 3 SLR**

Study Reference

Abbreviations: BW, body weight ; CDSR, Cochrane Database of Systematic Reviews; CINAHL, Cumulative Index to Nursing and Allied Health Literature; CMR, Cochrane Methodology Register; DARE, Database of Abstracts of Reviews of Effects; DIP, Diabetes in Pregnancy; GDM, gestational diabetes mellitus; HTA, Health Technology Assessment; IPD, individual patient data; LGA, large for gestational age; NHS EED, National Health Service Economic Evaluation Database; NR, not reported; OGCT, oral glucose challenge test; OGTT, oral glucose tolerance test; OR, odds ratio, SD, standard deviation.

Table 48: Beksac 2018

Study Reference	Beksac 2018
Study Design	Design Retrospective cohort study Objective To identify a cut-off value for the 50 g glucose challenge test (GCT) that predicts excess delivery weight. Dates January 2000 to December 2016 Country Turkey Setting Division of Perinatology, Department of Obstetrics and Gynecology, Hacettepe University Hospital, Ankara
Population Characteristics	Patient recruitment and eligibility **Recruitment:** Data were included from women who underwent GDM screening using the 50 g GCT at 24–28 weeks of pregnancy at the study institution **Inclusion criteria:** Women were singleton pregnancies who delivered live neonates after 28 weeks of pregnancy **Exclusion criteria:** Women with pregestational diabetes; women not screened with the 50 g GCT; women who required insulin therapy as a result of GDM screening were excluded to prevent a direct effect of insulin on the recorded delivery weight **Other:** The required data were obtained from the Hacettepe University Perinatal Medicine Database, which included information on referred high-risk pregnancies. Sample size N screened/invited = NR N eligible = NR N enrolled = NR N excluded (with reason) = NR N lost to follow-up = NR N completed = NR

N excluded from analysis = NR

Study Reference

Beksac 2018

N included in analysis = 584 Maternal demographics

	Maternal blood glucose level, mmol/L[a]			
Characteristic	**<7.770 (N=352)**	**7.770 to <8.880 (N=165)**	**8.880 to 9.990 (N=47)**	**>9.990 (N=20)**
Maternal age, years, median (range)	31 (18 to 45)	33 (21 to 44)	32 (21 to 42)	32 (24 to 37)
Obstetric history, n (%)				
Gravidity, median (range)	3 (1 to 13)	1 (0 to 5)	2 (1 to 8)	3 (1 to 6)
Parity, median (range)	1 (0 to 5)	1 (0 to 5)	1 (0 to 6)	1 (0 to 5)

[a] The 4 groups were defined using a 50 g glucose challenge test

Maternal glycaemic characteristics

Glucose tolerance	Study population (N=584)
50 g GCT, mmol/L, n (%)	
<7.770	352 (60.3)
7.770 to <8.880	165 (28.3)
8.880–9.990	47 (8.0)
>9.990	20 (3.4)

Duration of follow -up
NR

Method of blood glucose measurement
The 50 g GCT was performed at 24 to 28 weeks of pregnancy without any dietary restriction of carbohydrates before testing. A 50 g glucose load was administered orally, regardless of the time that had elapsed since the last meal. Venous plasma glucose levels were measured 1 hour later.

Diagnostic criteria and test(s)
- 50 g GCT
- Patients with 50 g GCT values of lower than 7.770 mmol/L were considered to be low risk and were not subjected to further testing. Women with values higher than 11.100 mmol/L were identified as having GDM. Those with values of 7.770–11.100 mmol/L underwent diagnostic testing with a 3 hour 100 g OGTT. A 100 g glucose load was administered orally in the morning to patients who had fasted overnight for a period of at least 8 hours.

Threshold cut-offs
According to the Carpenter–Coustan criteria, threshold values of 5.272 mmol/L for fasting glucose levels, 9.990 mmol/L for 1 hour levels, 8.602 mmol/L for 2 hour levels, and 7.770 mmol/L for 3 hour levels were used, with two or more values above the stated thresholds considered to be indicative of GDM. All of the patients diagnosed with GDM were counselled in the Endocrinology Division of Hacettepe University Hospital.

Methods Outcomes

Study Reference **Beksac 2018**

Pregnant women who met the inclusion criteria were stratified according to their 50 g GCT values: group 1 (<7.770 mmol/L); group 2 (7.770 to <8.880 mmol/L); group 3 (8.880–9.990 mmol/L), and group (>9.990 mmol/L). The following variables were assessed for each group: maternal age, gravidity, parity, pregnancy duration at delivery, 5 minute Apgar score and delivery weight.

Pregnancy outcomes

Outcome	<7.770 (N=352)	7.770 to <8.880 (N=165)	8.880 to 9.990 (N=47)	>9.990 (N=20)
Gestational age at birth, median (range)	37.0 (30 to 41)	37.0 (34 to 41)	38.0 (31 to 40)	37.5 (36 to 40)

Groups according to 50 g GCT

Neonatal outcomes according to maternal glycaemic status

Adverse neonatal outcomes	<table><tr><th>Outcome</th><th><7.770 (N=352)</th><th>7.770 to <8.880 (N=165)</th><th>8.880 to 9.990 (N=47)</th><th>>9.990 (N=20)</th></tr><tr><td>min Apgar score, median (range)</td><td>(6 to 10)</td><td>(6 to 10)</td><td>(10 to 10)</td><td>(9 to 10)</td></tr><tr><td>Delivery weight, g, median (range)</td><td>(1150 to 3910)</td><td>(1770 to 4150)</td><td>(2000 to 4280)</td><td>(2520 to 4320)</td></tr></table> Groups according to 50 g GCT Blood glucose (50 g GCT) was significantly associated with delivery weight in multiple linear regression using backward elimination method: unstandardized coefficients B=4.617 (SE 0.566); standardised coefficients β=0.287 (*t* 8.160) p<0.001
Study Reference	
Authors' Conclusions	The 50 g GCT can be used to identify women at risk of delivering offspring with excessive delivery weight.

Abbreviations: GCT: glucose challenge test; GDM: gestational diabetes; NR: not reported; SD: standard deviation.

Table 49: Berggren 2011

Study Reference	**Berggren 2011**
Study Design	Design Retrospective cohort study Objective To compare perinatal outcomes among women diagnosed with gestational diabetes by National Diabetes Data Group (NDDG) with women only meeting Carpenter-Coustan criteria Dates 1st April 1996 to 31st May 2010 Country US Setting UNC Women's Hospital, North Carolina
	Berggren 2011
	Patient recruitment and eligibility **Recruitment:** NR **Inclusion criteria:** Women eligible for GDM screening

Exclusion criteria: Women who delivered prior to 24 weeks' gestation, women with pre-gestational diabetes mellitus, and those without a documented GDM screening test result
Other: For multiple gestations, neonatal data for the firstborn were used
Sample size
N screened/invited = 41,398
N eligible = 33,179 N
enrolled = 5774
N excluded (with reason) = 320 (GDM diagnosis by 50 g 1 h glucose load results
N lost to follow-up = NR
N completed = 4659
N excluded from analysis = 0
N included in analysis = 4659 (GDM by CC only n=460; GDM by NDDG n=1082; Negative OGTT n=3117)

Population Characteristics

Maternal demographics

Characteristic	CC only (n=460)	Negative OGTT (n=3117)
Mean maternal age at delivery, years (SD)	30.6 (6.0)	29.4 (5.8)
Maternal age at delivery, n (%)		
≥ 35 years	113 (35)	559 (18)
< 35 years	347 (75)	2558 (82)
Pre-pregnant BMI, kg/m^2	NR	NR
BMI, kg/m^2	NR	NR
Weight, kg	NR	NR
Ethnicity, n (%)		
Caucasian	156 (34)	1215 (39)
African American	58 (13)	360 (12)
Latina	207 (45)	1338 (43)
Asian	29 (6)	162 (5)
Medical history, n (%)		
Chronic hypertension	39 (8)	138 (4)
History of pre-eclampsia	12 (2)	117 (4)
History of gestational diabetes	7 (2)	44 (1)
Prior C-section	77 (17)	537 (17)
Pre-pregnant smoking	NR	NR
Pre-pregnant alcohol use	NR	NR
Nulliparous, n (%)		

	Multiparity	304 (66)	1898 (61)
	Parous with GDM	NR	NR

Study Reference **Berggren 2011**

Education level	NR	NR

Maternal glycaemic characteristics

Glucose tolerance	CC only (n=460)	Negative OGTT (N=3117)
One-hour glucose load (mg/dL), median (IQR)	158 (149 to 173)	153 (145 to 163)

Study Reference

Methods

Duration of follow-up
Until delivery

Method of blood glucose measurement

GDM screening was performed between 24 and 28 weeks' gestation using a 50 g, 1-hour glucose load test

Diagnostic criteria and test(s)

Based on a 50 g, 1-hour glucose load test, plasma glucose values ≥140 mg/dL were considered screen-positive. Diagnostic testing was offered to these women and performed using a 100 g, 3-h OGTT. Women meeting National Diabetes Group (NDDG) criteria were diagnosed with GDM and received nutritional counselling and instruction for glucose self-monitoring. Women monitored capillary blood glucose with goals set as fasting < 105 mg/dL and 1 h postprandial <140 mg/dL or 2 h postprandial <130 mg/dL. Adequate glycaemic control at our institution was defined as 50% or more of blood glucose levels at goal levels. Medical therapy was initiated (subcutaneous insulin or oral glyburide) if adequate glycaemic control was not achieved with dietcontrol alone as determined by the primary obstetrical provider

Threshold cut-offs

The three study groups for this analysis included:

1) women who would be diagnosed with GDM only by CC criteria. Women who screened positive (1 h glucose load =140 mg/dL) but did not meet NDDG diagnostic criteria received routine prenatal care (CC only)

2) women diagnosed and treated for GDM by NDDG criteria, regardless of subsequent treatment (diet-control versus medical management with insulin or glyburide) required

3) women who screened positive but had a negative 3-hour oral glucose tolerance test and were not diagnosed with GDM by either criteria (negative OGTT)

Outcomes

Gestational age at delivery, preterm delivery (<37 weeks), mode of delivery (spontaneous vaginal delivery or C-section), 3rd or 4th degree perineal laceration, gestational hypertension, pre-eclampsia (composite of mild, severe eclampsia, and/or HELLP syndrome), birthweight (g), macrosomia (>4000 g), shoulder dystocia (abstracted from provider notation in perinatal record), NICU admission, and NICU stay >48 hours

Adverse neonatal outcomes

Pregnancy outcomes

Outcome, n (%) unless stated otherwise	CC only (n=460)	Negative OGTT (n=3117)	Adjusted Prevalence Ratio (95% CI)
Gestational age at birth, weeks, median (range)	39.3 (38.1 to 40.3)	39.3 (38.1 to 40.4)	NR
Pre-term birth	66 (14)	403 (13)	1.09 (0.86 to 1.39)
Pregnancy complications			
Gestational hypertension	33 (7)	150 (5)*	1.48 (1.02 to 2.13)
Berggren 2011			
Pre-eclampsia	58 (13)	264 (8)*	1.47 (1.02 to 2.13)
Stillbirth	NR	NR	NR
3rd/4th degree laceration	14 (3)	118 (4)	0.83 (0.48 to 1.44)

Prevalence ratios adjusted for controlling for parity, maternal delivery age over 35, ethnicity, and delivery year
*Significantly different

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)	CC only (n=460)	Negative OGTT (N=3117)	Adjusted Prevalence Ratio (95% CI)
Perinatal mortality	NR	NR	
Mode of birth			
Induction of labour	149 (32)	772 (25)*	NR
Normal spontaneous vaginal delivery	270 (59)	1923 (62)*	NR
Operative vaginal delivery			0.97 (0.68 to 1.39)
Vacuum-assisted vaginal delivery	11 (2)	141 (5)	NR
Forceps-assisted vaginal delivery	19 (4)	111 (4)	NR
C-section	160 (35)	942 (30)	1.16 (1.04 to 1.30)
Macrosomia	78 (17)	411 (13)*	1.25 (1.01 to 1.56)
LGA	NR	NR	NR
Birth injury			
Shoulder dystocia	24 (5)	109 (4)	1.41 (0.91 to 2.18)
Brachial plexus neuropathy	NR	NR	NR
Neonatal hypoglycaemia	NR	NR	NR

Admission to NICU	138 (30)	804 (26)	1.15 (0.99 to 1.33)
NICU stay over 48 hours	60 (43)	407 (52)	0.97 (0.76 to 1.25)
Long-term outcomes	NR	NR	NR

Prevalence ratios adjusted for controlling for parity, maternal delivery age over 35, ethnicity, and delivery year *Significantly different

Authors' Conclusions	Women who meet CC criteria but are not treated are at greater risk for hypertensive disorders of pregnancy and greater infant birthweight, compared to women diagnosed by NDDG and treated, as well as screen-positive women with a negative OGTT. These women who meet CC criteria, but not NDDG criteria, represent a group who would potentially benefit from treatment.

Abbreviations: BMI, body mass index; CC, Carpenter & Coustan; CI, confidence interval; GDM, gestational diabetes mellitus, HELLP, Hemolysis, Elevated Liver Enzymes, Low Platelets; LGA, large for gestational age; NDDG, National Diabetes Group; NICU, neonatal/newborn intensive care unit; NR, not reported; OGTT, oral glucose tolerance test; UNC, The University of North Carolina at Chapel Hill; US, United states.

Table 50: Berggren 2012 (MFMU Network)

Study Reference	Berggren 2012
Study Design	Design Secondary analysis of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units (MFMU) RCT for the treatment of mild GDM Objective To compare perinatal outcomes between self-identified Hispanic and non-Hispanic white women with mild GDM or glucose intolerance Dates NR Country US Setting NR
Characteristics	Patient recruitment and eligibility **Recruitment:** Secondary analysis of RCT data: in initial randomisation, eligible women with an elevated 50 g 1 h glucose load result underwent a 100 g 3 h OGTT, and these results diagnosed mild GDM. Women were randomised to treatment vs no treatment, matched for race/ethnicity and body mass index (BMI < vs > 27 mg/kg^2). Women with normal OGTT results were enrolled as the observational cohort. **Inclusion criteria:** Women with a 1 hour glucose load test result between 135 and 200 mg/dL between 12 0/7 weeks and 30 6/7 weeks' gestation **Exclusion criteria:** Women with pre-gestational diabetes; an abnormal glucose screening test prior to 24 weeks' gestation; a history of GDM, stillbirth, multi-fetal gestation, asthma or chronic hypertension; if taking corticosteroids; or if imminent preterm delivery was anticipated Sample size N screened/invited = NA N eligible = 1889 N enrolled = 1535 N excluded (with reason) = NR **Population** N lost to follow-up = NR

Characteristic	Glucose intolerant (n=767)			Mild untreated GDM (n=371)		
	Hispanic	Non-Hispanic White	p-value	Hispanic	Non-Hispanic White	p-value
Mean maternal age, years (SD)	27.6 (5.4)	28.0 (5.4)	0.29	29.5 (5.6)	28.5 (5.0)	0.08
Mean gestational age, weeks (SD)	28.4 (1.5)	29.2 (1.3)	<0.001	28.6 (1.5)	29.5 (1.3)	<0.001

Study Reference

N completed = NR
N excluded from analysis = NR
N included in analysis = 1535 (Hispanic or non-Hispanic White)

Maternal demographics

Berggren 2012

Cardiometabolic health						
BMI, kg/m²	30.1 (4.5)	29.5 (5.3)	0.11	30.2 (4.3)	30.6 (6.2)	0.51
Medical history/risk factors, n (%)						
Hypertension	NR	NR	NR	NR	NR	NR
Diabetes	NR	NR	NR	NR	NR	NR
Pre-pregnant smoking, n (%)	16 (3.1)	48 (19.4)	<0.001	4 (1.6)	17 (15.0)	<0.001
Pre-pregnant alcohol use	NR	NR	NR	NR	NR	NR
Obstetric history, n (%)						
Primigravida, mean (SD)	136 (26.2)	107 (43.3)	<0.001	75 (29.4)	48 (41.4)	0.02
Education level	NR	NR	NR	NR	NR	NR

Maternal glycaemic characteristics

Glucose tolerance	Glucose intolerant (n=767)			Mild untreated GDM (n=371)		
	Hispanic	Non-Hispanic White	p-value	Hispanic	Non-Hispanic White	P-value
50 g 1 hour oral glucose load (mg/dL)	152.6 (13.1)	153.1 (13.3)	0.57	160.6 (15.5)	159.5 (15.9)	0.51
100 g 3-hour oral glucose tolerance test (mg/dL)						
Fasting	84.7 (5.8)	85.0 (5.8)	0.50	86.3 (5.8)	86.3 (5.6)	0.90
1 hour	156.3 (23.4)	151.2 (26.0)	0.006	193.8 (18.3)	192.1 (21.9)	0.46
2 hours	130.1 (22.0)	130.5 (21.6)	0.82	172.5 (21.1)	172.6 (16.4)	0.94
3 hours	111.6 (21.0)	105.3 (23.2)	0.0002	136.7 (29.2)	128.6 (32.2)	0.02

Study Reference

Methods

Duration of follow-up
NR

Method of blood glucose measurement
50 g 1 hour screening test, 100 g 3 hour OGTT

Diagnostic criteria and test(s)
In initial randomisation, eligible women with an elevated 50g 1-hr glucose load result underwent a 100g 3-hr OGTT, and these results diagnosed mild GDM. Women with normal OGTT results were enrolled as the observational cohort

Threshold cut-offs
Women were classified into one of 3 groups in the parent study:

- 767 women with glucose intolerance with an elevated 50g 1 hour screening test (≥135, but <200 mg/dL) but normal 3 hour OGTT, matched to the randomised patients by clinical centre, race/ethnicity, and BMI

Berggren 2012

- 371 women with mild GDM as diagnosed by a fasting glucose <95 mg/dL but two or more 100g 3-hour oral glucose tolerance test (OGTT) results at or above established thresholds who were randomised to no treatment
- 397 women also diagnosed with mild GDM by the same criteria as above but who were randomised to treatment.

Outcomes
Primary endpoint
Composite adverse perinatal outcome for the original study was occurrence of one or more of the following: perinatal mortality (stillbirth or perinatal death), hypoglycaemia, hyperbilirubinemia, elevated cord blood C-peptide level, or birth trauma. overall and individual components of the composite outcome: hyperbilirubinemia, elevated cord blood C-peptide, and hypoglycaemia. Perinatal mortality and birth trauma are not included in individual analyses due to small numbers for each.

Secondary endpoints

Gestational age at delivery (weeks' gestation, preterm birth <37 weeks), birthweight (grams, small for gestational age, large for gestational age >90th centile, macrosomia >4000 g), either gestational hypertension or pre-eclampsia, and neonatal intensive care unit (NICU) admission.

Pregnancy outcomes

Outcome	Glucose intolerant (n=767)			M ld untreated GDM (n=371)		
	Hispanic	Non-Hispanic white	aOR (95% CI) or β coefficient (SE)	Hispanic	Non-Hispanic white	aOR (95% CI) or B coefficient (SE)
Gestational age at birth, weeks, mean (SD)	39.4 (1.6)	39.1 (1.5)	0.39 (0.14), p=0.005	39.2 (1.6)	38.7 (1.9)	0.48 (0.21), p=0.02

	Pre-term birth (before 37 weeks), n (%)	35 (7)	14 (6)	1.58 (0.75 to 3.36)	23 (9)	14 (12)	0.61 (0.28 to 1.33)

Study Reference

Pregnancy complications						
Gestational hypertension or preeclampsia	38 (7)	27 (11)	0.73 (0.41 to 1.30)	37 (15)	13 (11)	1.71 (0.78 to 3.71)

Adverse neonatal outcomes

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)	Glucose intolerant (n=767)			Mild untreated GDM (n=371)		
	Hispanic	Non-Hispanic white	aOR (95% CI) or B coefficient (SE), p-value	Hispanic	Non-Hispanic white	aOR (95% CI) or B coefficient (SE)
Perinatal mortality	NR	NR	NR	NR	NR	NR
Mode of birth	NR	NR	NR	NR	NR	NR
Birthweight, g, mean (SD)	3431 (499)	3344 (510)	31.7 (41.9), p=0.45)	3478 (543)	3388 (630)	34.0 (69.1), p=0.62)
Macrosomia, n (%)	62 (12)	23 (9)	1.12 (0.63 to 1.98)	40 (16)	17 (15)	1.01 (0.52 to 1.96)
LGA, n (%)	63 (12)	22 (9)	1.19 (0.67 to 2.11)	38 (15)	16 (14)	0.94 (0.47 to 1.86)
Birth injury	NR	NR	NR	NR	NR	NR
Neonatal hypoglycaemia, n (%)	84 (21)	25 (13)	2.04 (1.18 to 3.53)	30 (15)	13 (14)	0.98 (0.44 to 2.18)

Study Reference	Berggren 2012					
Admission to NICU, n (%)	30 (6)	19 (8)	0.97 (0.48 to 1.94)	21 (8)	13 (11)	0.63 (0.28 to 1.41)
Long-term outcomes	NR	NR	NR	NR	NR	NR

Authors' Conclusions	Additional efforts may target at-risk women with hyperglycaemia, but not overt GDM, for intervention and treatment, regardless of race/ethnicity. Our findings suggest that diagnostic criteria tailored to race/ethnicity may not be warranted, at least not among women with mild GDM or glucose intolerance.

Abbreviations: aOR, adjusted odds ratio; BMI, body mass index; CI, confidence interval; GDM, gestational diabetes mellitus, MFMU, maternal-fetal medicine units; NICHD, National Institute of Child Health and Human Development; NICU, neonatal/newborn intensive care unit; NR, not reported; OGTT, oral glucose tolerance test; RCT, randomised control trial; SD, standard deviation; SE, standard error; US, United states.

Table 51: MAMMA, Berntorp 2015

Study Reference	MAMMA, Berntorp 2015
Study Design	Design Prospective cohort study Objective To evaluate the relative importance of maternal BMI and glucose levels in prediction of large-for-gestational-age (LGA) births Dates 2003 to 2005 Country Sweden Setting Delivery departments, Skåne, southern Sweden
Population	Patient recruitment and eligibility **Recruitment:** During the years 2003–2005, pregnant women representing different glucose categories according to universal screening by the 2 h glucose level of the OGTT were invited to take part in a follow-up program, the Mamma Study. During the recruitment period, OGTT results from the local antenatal clinics were sent to the study co-ordinator, enabling the identification of the test results of women who consented to be enrolled. If a woman had repeated pregnancies during the period, only the first one was included. If a repeat OGTT was performed, only the first one was included. **Inclusion criteria:** NR **Exclusion criteria:** NR **Characteristics** **Other:** Sample size

N screened/invited =

Study Reference

MAMMA, Berntorp 2015

N eligible = 11,976 OGTT results N enrolled
= NR
N excluded (with reason) = 10 974 (missing information from perinatal database or on LGA excluded)
N lost to follow-up = NR N
completed = NR
N excluded from analysis = NR
N included in analysis =development sample 5487; validation sample 5487

Maternal demographics

Characteristic	Glucose quartiles (mmol/L)		
	<5.7 (n=2637)	5.7 to 6.4 (n=2783)	6.5 to 7.2 (n=2,819)
Maternal age, years			
<20	80 (32.5)	62 (25.2)	63 (25.6)
20–34	2148 (24.2)	2288 (25.8)	2264 (22.5)
≥35	409 (21.6)	433 (22.9)	492 (26.0)
Cardiometabolic health			
Maternal BMI, kg/m^2			
<18.5	50 (25.6)	50 (25.6)	50 (25.6)
18.5 to 24	1496 (25.1)	1569 (26.3)	1542 (25.9)
25.0 to 29.9	585 (22.0)	641 (24.1)	687 (25.9)
30 to 34.9	182 (20.8)	187 (21.4)	223 (25.5)
≥35	83 (20.1)	103 (25.0)	93 (22.6)
Weight, kg	NR	NR	NR
Ethnicity, n (%)			
White	NR	NR	NR
Black	NR	NR	NR
South Asian	NR	NR	NR
East Asian	NR	NR	NR
Mixed	NR	NR	NR

	Medical history/risk factors, n (%)			
	Hypertension	NR	NR	NR
	Diabetes	NR	NR	NR
Study Reference	**MAMMA, Berntorp 2015**			
	Pre-pregnant smoking	NR	NR	NR
	No	2220 (23.4)	2408 (25.4)	2430 (25.6)
	Yes	341 (27.2)	309 (24.6)	333 (26.6)
	Obstetric history, n) (%			
	Parity			
	1	128 (23.8)	134 (24.9)	141 (26.2)
	2 to 3	119 (24.1)	128 (26.0)	124 (25.2)
	≥4	16 (24.1)	15 (22.5)	15 (23.4)
	Education level	NR	NR	NR

Maternal glycaemic characteristics

NR

Study Reference	MAMMA, Berntorp 2015
Methods	Duration of follow-up NR Method of blood glucose measurement The HemoCue blood glucose system was used to obtain immediate analysis of glucose concentrations. All women received a 75 g OGTT. Diagnostic criteria and test(s) The 75 g OGTT is offered to all women in the 28th week of gestation and is done after overnight fasting at their local antenatal clinic. Normal glucose tolerance during pregnancy is defined as a 2 h capillary plasma glucose concentration < 8.9 mmol/L. In 2004, routine glucose measurements in Sweden were switched from blood glucose measurements to plasma glucose measurements, and a transformation factor of 1.11 was agreed on, resulting in a 2 h threshold value of 10.0 mmol/L for capillary plasma glucose to define GDM. If 2 h capillary plasma glucose concentration is 8.9–9.9 mmol/L, indicating gestational impaired glucose tolerance (IGT), the OGTT is repeated within a week. Threshold cut-offs NGT: <8.9 mmol/L IGT: 8.9 to 9.9 mmol/L GDM: 10.0 mmol/L Outcomes LGA births, small-for-gestational-age (SGA) births and adequate-for-gestational-age (AGA) births were defined as birth weight greater than +2 standard deviations (SD), less than −2 SD and between −2 SD and +2 SD of the expected birth weight for gestational age and gender, respectively, according to the Swedish reference curve for fetal growth. The prediction model for LGA was developed on the development dataset using univariate and multivariable logistic regression analyses. The variables tested were: maternal age, parity 1, parity ≥ 4 (with parity 2–3 as reference), maternal smoking, maternal BMI (continuous), maternal height (continuous), and glucose levels (continuous). Variables with a crude p-value of <0.05 in their association with LGA in the univariate model were entered into a multiple model, and variables with a p-value of <0.05 in the multiple model were entered into the final multiple model. A two-sided p-value <0.05 was considered statistically significant.
Adverse neonatal	Pregnancy outcomes

Outcome, n (%)*	<5.7 (n=2637)	5.7 to 6.4 (n=2783)	6.5 to 7.2 (n=2819)	p-value
Gestational age at birth				
<37 weeks (i.e. pre-term)	117 (20.0)	148 (25.3)	153 (26.2)	0.006
37–41+6 weeks	2345 (24.0)	2472 (25.3)	2502 (25.6)	
≥42+0 weeks	175 (26.5)	163 (24.7)	164 (24.8)	

* % is out of total in row rather than total in glucose group

Study Reference **MAMMA, Berntorp 2015**

outcomes

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)*	<5.7 (n=2637)	5.7 to 6.4 (n=2783)	6.5 to 7.2 (n=2819)	p-value
Perinatal mortality	NR	NR	NR	NR
Mode of birth	NR	NR	NR	NR
Macrosomia	NR	NR	NR	NR
Weight for gestational age				

Study Reference	MAMMA, Berntorp 2015			
SGA	69 (23.2)	80 (26.9)	68 (22.9)	<0.001
AGA	2446 (24.2)	2577 (25.5)	2578 (25.6)	
LGA	115 (20.1)	110 (19.2)	156 (27.3)	
Birth injury	NR	NR	NR	NR
Neonatal hypoglycaemia	NR	NR	NR	NR
Admission to NICU	NR	NR	NR	NR
Long-term outcomes	NR	NR	NR	NR

* % is out of total in row rather than total in glucose group

OR for LGA per mmol/L of glucose or per SD of glucose

Risk factor	Univariate model		Multiple model		Final multiple model		
	OR	p-value	OR	p-value	OR	95% CI	p-value
2-hour glucose (per 1 mmol increase)	1.12	0.003	1.09	0.033	1.09	1.01 to 1.18	0.028

Authors' Conclusions	Both the 2 h glucose level of the OGTT and maternal BMI had a significant effect on the risk of delivering an LGA neonate. However, the relative contribution was much higher for BMI, even when taking other risk factors into account.

Abbreviations: AGA, adequate-for-gestational-age; BMI, body mass index; CI, confidence interval; GDM, gestational diabetes mellitus, IGT, impaired glucose tolerance; LGA, large-for-gestational-age; NGT, normal glucose tolerance; NICU, neonatal/newborn intensive care unit; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; SD, standard deviation; SGA, small-for-gestational-age.

Table 52: Biri 2009

Study Reference	Biri 2009
Study Reference	

Study Design

Design
Retrospective analysis

Objective
To evaluate the effect of markedly elevated 50-g glucose loading test (GLT) (≥200 mg/dL) and equivocal 100-g GLT (one abnormal value) results on maternal and perinatal outcomes.

Dates

January 2004 to December 2006
Country
Turkey
Setting

Faculty of Medicine, Department of Obstetrics and Gynecology, Gazi University

Biri 2009

Patient recruitment and eligibility

Recruitment: Retrospective analysis of all singleton pregnancies that were screened for GDM in the study institution **Inclusion criteria:** NR **Exclusion criteria:** NR

Other: NR Sample size

Population Characteristics

N screened/invited = 2029
N eligible = 2029
N enrolled = 2029
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = 2029
N excluded from analysis = 0 N included in analysis = 2029

Maternal demographics

Characteristic	Group 1 (N=1432)	Group 2 (N=326)	Group 3 (N=142)

Mean maternal age, years (SD)	29.6 (4.6)	30.9 (4.9)	32.1 (4.6)
Cardiometabolic health	NR	NR	NR
Ethnicity, n (%)	NR	NR	NR
Medical history/risk factors, n (%)	NR	NR	NR
Obstetric history, n (%)	NR	NR	NR
Education level	NR	NR	NR

Study Reference	
	Maternal glycaemic characteristics NR
Methods	Duration of follow-up To delivery Method of blood glucose measurement GLT (50 g) was performed for 2059 patients between the 24th and 28th gestational weeks as recommended by ACOG Diagnostic criteria and test(s) GLT (50 g): a value of 140 mg/dL was considered as the cut-off. Patients who were screen-positive underwent a 100 g GTT to diagnose GDM as 100 g GTT was preferred to 75 g GTT in the authors' department. Cut-off values for plasma glucose were defined as 105, 190, 165, 145 mg/dL for fasting, 1, 2 and 3 h tests after the 100 g GTT, respectively. These cut-off values adopted by the study department were first proposed by O'Sullivan and Mahan in 1964 and were converted to plasma values by the 'National Diabetes Data Group' in 1979 Threshold cut-offs

Biri 2009

The first group consisted of patients with a normal 50 g GLT. Second group was formed by patients with an abnormal 50 g but a normal 100 g GLT. Third group included patients with one abnormal value after 100 g GLT. Patients in the fourth group were diagnosed to have GDM after an abnormal 100 g GLT. Patients in the fifth group had a value ≥200 mg/dL after 50 g GLT and were diagnosed to have GDM.

- 1432 patients (70.6%) had a value below 140 mg/dL after 50 g GLT and formed group 1
- Group 2 consisted of the 326 patients (16.1%) with an abnormal 50 g GLT and a normal 100 g GTT
- 142 patients (7.0%) with an abnormal 50 g GLT and only one abnormal value detected by 100 g GTT fell into group 3
- Group 4 consisted of 73 patients (3.6%) with two or more abnormal values detected in 100 g GTT performed after a GLT value between 140 mg/dL and 199 mg/dL
- 56 patients (2.8%) with a GLT value ≥200 mg/dL for whom GTT was not performed formed group 5

Outcomes

Maternal ages, gestational ages at birth, birth weights, Apgar scores and neonatal complications were the main parameters studied. Neonates with a birth weight below the 10th percentile were defined as small for gestational age (SGA) and those with a birth weight above the 90th percentile were defined as large for gestational age (LGA). A cut-off value of 4000 g was considered for the definition of macrosomia. Blood glucose levels of neonates were evaluated 1 and 4 h after birth and venous haematocrit levels were evaluated 4 h after birth. Hypoglycaemia was defined as a blood glucose level below 40 mg/dL and polycythaemia was defined as a venous haematocrit level above 65%

Pregnancy outcomes

Outcome, %	Group 1 (n=1432)	Group 2 (n=326)	Group 3 (n=142)
Gestational age at birth, mean (SD)	39.0 (1.4)	38.6 (1.3)	38.4 (1.4)
Pre-term birth, %	0.4	0.6	1.4

Study Reference

Pregnancy complications, %			
Pre-eclampsia	1.5	2.3	2.1
Stillbirth	NR	NR	NR

Neonatal outcomes according to maternal glycaemic status

Adverse neonatal outcomes	**Outcome, %**	**Group 1 (n=1432)**	**Group 2 (n=326)**	**Group 3 (n=142)**
	Perinatal mortality	NR	NR	NR
	C-section	54.8	63.1	63.4
	Macrosomia	5.8	8.3	12.7
	LGA	8.0	12.0	14.8
	Birth injury	NR	NR	NR
	Neonatal hypoglycaemia	0.4	1.2	3.5
	Neonatal hospitalisation	5.9	9.7	14.8
	Mean 1-min Apgar score (SD)	9.1 (0.8)	9.0 (0.8)	9.0 (0.8)
	Mean 5-min Apgar score (SD)	9.9 (0.4)	9.9 (0.4)	9.9 (0.4)
	Long-term outcomes	NR	NR	NR
Study Reference	**Biri 2009**			
Authors' Conclusions	Adverse maternal and perinatal outcomes in patients with one elevated GTT value and in patients with a GLT value >4200 mg/dL detected in this study warrant close glucose monitoring and treatment in these groups even in the absence of a diagnostic abnormal GTT			

Abbreviations: ACOG, American College of Obstetricians and Gynecologists; GDM, gestational diabetes mellitus; GLT, glucose loading test; GTT, glucose tolerance test; LGA, large-for-gestational-age; NR, not reported; SD, standard deviation; SGA, small-for-gestational-age.

Table 53: Cheng 2009

Study Reference	**Cheng 2009**
Study Design	Design Retrospective cohort study Objective To examine perinatal outcomes in women who would meet the diagnostic criteria for gestational diabetes mellitus (GDM) according to the Carpenter and Coustan but not according to the National Diabetes Data Group (NDDG) thresholds. Dates January 1988 to December 2001 Country US Setting University of California, San Francisco

	Patient recruitment and eligibility **Recruitment:** A retrospective cohort study of all pregnancies screened for GDM and delivered at the University of California, San Francisco **Inclusion criteria:** NR **Exclusion criteria:** The exclusion criteria were multifetal pregnancies, vaginal breech deliveries, delivery before 24 weeks of gestation, known lethal congenital anomalies, and pregestational diabetes mellitus. **Other:** NR Sample size
Population Characteristics	N screened/invited = NR N eligible = 14,693 N enrolled = NR N excluded (with reason) = NR N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = NR

Study Reference **Cheng 2009**

Maternal demographics

Characteristic	No GDM (n=13,940)	GDM only by C&C (untreated; n=273)
Maternal age, years, %		
	82.1	72.9
	17.9	27.1
-pregnant BMI, [2]	NR	NR
	10.2	13.9
	89.8	86.1
	NR	*NR*
<35 (n=11,966)		27.0
≥35 (n=2,727)		6.0
Pre kg/m		12.0
BMI, kg/m²		55.0
< 30 (n=8,172)		
≥30 (n=969)		NR
Weight, kg		NR
Ethnicity, %		NR
White (n=5,316)		
African American (n=2,111)		
Latina (n=1,774)		
Asian (n=4,773)		
Medical history		
Hypertension		
Diabetes		
	38.9	
	15.4	

Study Reference **Cheng 2009**

	12.6	
	33.1	
	NR	
	NR	
Pre-pregnant smoking	NR	
Pre-pregnant alcohol use	NR	NR
Nulliparous (n=7,938)	54.4	48.0
Multiparous (n=6,755)	45.6	52.0
Education level	NR	NR

Methods

Maternal glycaemic characteristics NR Duration of follow-up
Until delivery

Method of blood glucose measurement
Screening of GDM was most often performed between 24 and 28 weeks of gestation using the 50 g, 1 h screening test, with subsequent 3 h 100 g OGTT for confirmation if screened positive. During the study period, plasma glucose was measured by the glucose oxidase technique. The test was switched to the glucose hexokinase technique between 1992 and 1999. However, internal controls were used to confirm consistency between the two techniques, and the equipment was calibrated three times daily for quality control. Test results were abstracted from a laboratory database and linked with a perinatal database

Diagnostic criteria and test(s)
In women at high risk for GDM, early screening and diagnosis during the first or early second trimester was commenced.
During the study period, GDM was diagnosed using the NDDG criteria at the University of California, San Francisco. Women who would have been given a diagnosis of GDM based on the Carpenter and Coustan (but not the NDDG) criteria received routine care and did not receive further counselling or nutrition education during the study period, because they were considered “ruled out” for GDM

Threshold cut-offs
In these high-risk patients who screened negative, a repeat 50 g, 1 h screening test was performed between 24 and 28 weeks of gestation. For the majority of the study population (more than 98%), the threshold for obtaining an OGTT was 140 mg/dL since 1995. Before this, a small minority (1–3%) of the women were considered screened positive when a 50 g, 1 h screening threshold of 135 mg/dL was used

Outcomes
Maternal outcomes included mode of delivery, third- or fourth- degree perineal lacerations, pre-eclampsia, and preterm delivery (less than 37 weeks of gestation)

Study Reference **Cheng 2009**

Neonatal outcomes included 5-minute Apgar score <7, neonatal acidaemia as measured by umbilical cord artery pH <7.0 and base excess <–12, LGA (defined as birth weight >97th centile by gestational age), macrosomia (birth weight >4,500 g), shoulder dystocia (as diagnosed by the delivering attending), birth trauma (examined as a composite variable for brachial plexus injury, facial nerve palsy, clavicular fracture, skull fracture, and head laceration), neonatal hypoglycaemia, jaundice, and admissions to the intensive care nursery

Pregnancy outcomes

Outcome, n (%)	No GDM (n=13,940)	GDM by C&C Only (n=273)	AOR (95% CI)	P value
Gestational age at birth	NR	NR	NR	NR
Pre-term birth (n=1,057)	7.0	9.5	1.36 (0.84 to 2.18)	0.09
Pregnancy complications				
Pre-eclampsia (n=677)	4.5	6.2	1.30 (0.71 to 2.38)	0.01
Stillbirth	NR	NR	NR	NR
3rd/4th degree perineal laceration (n=1,108)	9.0	11.4	1.16 (0.73 to 1.86)	0.14
Postpartum haemorrhage (n=3,297)	22.6	26.7	1.08 (0.79 to 1.49)	<0.001

Study Reference **Cheng 2009**

Neonatal outcomes according to maternal glycaemic status

Adverse neonatal outcomes

Outcome, n (%)	No GDM (n=13,940)	GDM by C&C Only (n=273)	AOR (95% CI)	P value
Perinatal mortality	NR	NR	NR	NR
Mode of birth				
Spontaneous vaginal delivery (n=10,125)	69.1	58.5	-	<0.001
Operative vaginal delivery (n=2,053)	14.0	18.8	1.72 (1.20 to 2.46)	
C-section (n=2,525)	16.9	22.7	1.44 (1.01 to 2.07)	
Labour dystocia	43.0	42.9	1.37 (0.82 to 2.28)	0.03
Fetal intolerance of labour	21.1	10.2	0.66 (0.24 to 1.80)	
Repeat	14.4	14.3	-	
Other	21.5	32.6	-	
Macrosomia (n=235)	1.6	4.0	4.47 (2.26 to 8.86)	0.01
LGA (n=209)	1.3	5.1	4.28 (2.24 to 8.18)	<0.001
5-min Apgar score <7 (n=388)	2.6	2.6	1.01 (0.48 to 2.53)	0.88
Birth trauma[a] (n=542)	3.7	4.4	1.26 (0.66 to 2.42)	0.43
Shoulder dystocia (n=209)	1.7	3.3	2.24 (1.03 to 4.88)	0.14
Neonatal hypoglycaemia (n=269)	1.7	1.8	0.93 (0.34 to 2.55)	<0.001
Admission to Intensive Care Nursery Admission (N=884)	6.0	5.9	0.99 (0.54 to 1.77)	0.91
Long-term outcomes	NR	NR	NR	NR

[a] Composite variable for skull fractures, head lacerations, clavicular fractures, facial nerve palsy, and Erb's palsy.

Study Reference	Cheng 2009
Authors' Conclusions	Women diagnosed with GDM by the Carpenter and Coustan criteria but not by the NDDG criteria had higher risk of operative deliveries, macrosomia, and shoulder dystocia. The authors recommend using the Carpenter and Coustan diagnostic thresholds for GDM, because these diagnostic criteria are more sensitive than the NDDG criteria.

Abbreviations: AOR, adjusted odds ratio; C&C, Carpenter and Coustan; CI, confidence interval; GDM, gestational diabetes mellitus, LGA, large-for-gestational-age; NDDG, National Diabetes Data Group; NR, not reported; OGTT, oral glucose tolerance test; US, United states.

Table 54: Corrado 2009

Study Reference	Corrado 2009
Study Design	Design Retrospective cohort study Objective To evaluate which pregnant women with a single abnormal value in the oral glucose tolerance test are at increased risk for adverse perinatal outcome. Dates January 1996 to December 2005 Country Italy Setting Department of Obstetrics and Gynecology, University of Messina
Population Characteristics	Patient recruitment and eligibility **Recruitment:** Retrospective enrolment **Inclusion criteria:** Caucasian singleton pregnancies who had a positive screening test and then an OGTT during the study period **Exclusion criteria:** Multiple gestations were excluded from the study **Other:** Sample size N screened/invited = 989 N eligible = 989 N enrolled = 776 N excluded (with reason) = 142 with GDM N lost to follow-up = NR N completed = NR N excluded from analysis = 71 (missing obstetric outcome data)

Study Reference **Corrado 2009**

N included in analysis =776

Maternal demographics

Characteristic	OAV (n=152)	Controls (n=624)	Significance
Mean age, years (SD)	31.2 (5.06)	30.10 (4.85)	0.01
Cardiometabolic health, mean (SD)			
Pre-pregnant BMI, kg/m^2	NR	NR	NR
BMI, kg/m^2	25.01 (5.14)	24.15 (4.37)	0.04
Weight, kg	NR	NR	NR
Ethnicity, n (%)	NR	NR	NR
Medical history/risk factors, n (%)			
Hypertension	NR	NR	NR
Diabetes	NR	NR	NR
Family history of diabetes	54 (35.5)	173 (27.7)	0.06
Pre-pregnant smoking	NR	NR	NR
Pre-pregnant alcohol use	NR	NR	NR
Obstetric history, n (%)			
Nulliparous	NR	NR	NR
Parity >1	66 (43.4)	276 (44.1)	0.92
Education level	NR	NR	NR

Maternal glycaemic characteristics NR

Study Reference

Methods

Duration of follow-up
Data of the maternal-fetal outcome and weight gain in pregnancy were collected after delivery from patients' charts

Method of blood glucose measurement
The pregnant women were routinely screened at 24–28 weeks of gestation. The glucose oxidase method was used for plasma venous glucose determination.

Diagnostic criteria and test(s)
/enous plasma glucose concentrations were measured 1 h after a 50 g oral glucose load (GCT). If the glucose value was ≥135 mg/dL the subject underwent a 3 h–100 g OGTT within the next 2 weeks. GDM was diagnosed when two or more glucose values equalled or exceeded 95, 180, 155 and 140 mg/dL, respectively, according to Carpenter's criteria. The plasma insulin concentration was routinely measured on the fasting blood sample.

Threshold cut-offs
GDM: two or more glucose values ≥95, 180, 155 and 140 mg/dL respectively according to Carpenter's criteria
Patients were divided into two groups: one abnormal value (OAV) and control patients (with all four glucose values within the normal range at the OGTT).

Outcomes
Primary endpoint

Corrado 2009

- Hypertensive disorders of pregnancy (pre-eclampsia and pregnancy-induced hypertension), caesarean section, gestational age at delivery, birth weight, macrosomia, Apgar score at 1 and 5 minutes after birth and neonatal hypoglycaemia
- Pre-eclampsia was defined as blood pressure higher than 140/90 on two or more occasions and proteinuria >300 mg in 24 h. Pregnancy induced hypertension was diagnosed if the blood pressure met the previously mentioned criteria without the presence of proteinuria. Macrosomia was defined as a birth weight ≥4000 g and hypoglycaemia a glucose value 530 mg/dl within 2 h from the birth.

Secondary endpoints

Pregnancy outcomes

Outcome	OAV (n=152)	Controls (n=624)	p-value
Gestational age at birth, mean (SD)	38.5 (1.8)	38.8 (1.6)	0.06
Pre-term birth	NR	NR	NR
Pregnancy complications, n (%)			
Hypertensive disorders	21 (13.8)	27 (4.3)	0.0001

Neonatal outcomes according to maternal glycaemic status

	Outcome	OAV (n=152)	Controls (n=624)	p-value
Adverse neonatal outcomes	**Perinatal mortality**	NR	NR	NR
	Mode of birth, n (%)			
	Induction of labour	NR	NR	NR
	Vaginal delivery	NR	NR	NR
	Instrumental delivery	NR	NR	NR
	Caesarean sections, n (%)	(56)	(39)	0.0001
	Planned C-section	NR	NR	NR
	Macrosomia	(12.5)	(6.2)	0.01
	LGA	NR	NR	NR
	Birth injury	NR	NR	NR
	Apgar Score, mean (SD)			
	1-minute score	7.9 (1.9)	8.1 (1.7)	0.3
	5-minute score	9.3 (0.9)	9.4 (0.7)	0.2
	Neonatal hypoglycaemia, n (%)	(6.2)	(4.1)	0.4
	Admission to NICU	NR	NR	NR
	Long-term outcomes	NR	NR	NR
Authors' Conclusions	The authors' results show that the implications of a single elevated glucose tolerance test value vary in relation to the timing of the abnormal value. In fact, OAV fasting or 1-h after load has a higher prevalence for an adverse obstetric outcome, whereas a 2 or 3-h value does not present significant differences when compared with the control group.			

Abbreviations: BMI, body mass index; GCT, glucose challenge test; GDM, gestational diabetes mellitus; LGA, large-for-gestational-age; NICU, neonatal/newborn intensive care unit; NR, not reported; OAV, one abnormal value; OGTT, oral glucose tolerance test; SD, standard deviation.

Table 55: Davis 2018

Study Reference **Davis 2018**

Study Design

Design
Retrospective cohort study

Objective
To examine the association between different diagnostic criteria for GDM and adverse birth outcomes

Dates
January 2006 to December 2010

Country
US

Setting
A large women's academic hospital

Study Reference

Population

Patient recruitment and eligibility

Recruitment: Retrospective study of pregnant women with singleton pregnancies who delivered at the study institution.

Inclusion criteria: Participants had to have a 1 h 50 g GCT <130 mg/dL, or ≥130 mg/dL and <180 mg/dL, and a clinically indicated 3 h 100 g OGTT

Exclusion criteria: Women were excluded if they had multiple gestations, pre-existing diabetes, or delivered at another hospital. Women were also excluded if they were missing key independent variables, such as glucose values or date of last period, had out of range GAs (<0 or >43 weeks) or did not have glucose testing done

Sample size
N screened/invited = 7,819
N eligible = 7,819
N enrolled = 6,894
N excluded (with reason) = missing date of last period: 554; did not have GDM testing with 50 or 100 g OGTT: 91; gestational age at testing was out of range (<0 or >43 weeks): 42; gestational age at delivery was out of range (<0 or >43 weeks): 192; no glucose testing was done: 46 **Characteristics**
N lost to follow-up = NR
N completed = 5,937
N excluded from analysis = 755 (GDM test data not useable e.g. only a 100 g test was done, or 50 g test only and test result was 136–179 [inclusive], or GDM testing pattern was not "50 GCT or 50 + 100 g OGTT) N
included in analysis = 5,937

Maternal demographics

	Normal (n=4,941)	Elevated GCT + NL OGTT (n=544)	GDM by IADPSG criteria (n=181)
Maternal age at delivery, years, mean (SD)	30.3 (5.7)	31.9 (5.1)	32.1 (5.2)
Cardiometabolic health, n (%)			

Study Reference **Davis 2018**

Pre-pregnant BMI, kg/m^2	NR	NR	NR
Underweight	155 (4.4)	10 (2.9)	1 (0.9)
Normal weight	2,124 (59.9)	204 (58.5)	58 (52.3)
Overweight	759 (21.4)	72 (20.6)	33 (29.7)
Obese	509 (14.4)	63 (18.1)	19 (17.1)
Ethnicity, n (%)			
White	3,500 (70.8)	409 (75.2)	134 (74.0)
Black	942 (19.1)	51 (9.4)	23 (12.7)
Other	348 (7.0)	65 (11.9)	17 (9.4)
Unknown	151 (3.1)	19 (3.5)	7 (3.9)
Medical history/risk factors, n (%)			
Any smoking during pregnancy	367 (7.4)	30 (5.5)	15 (8.3)
Obstetric history, n (%)			
Nulliparous	2,150 (43.5)	251 (46.2)	82 (45.3)
Primiparous	1,880 (38.1)	202 (37.2)	65 (35.9)
Multiparous (2+ or more)	907 (18.4)	90 (16.6)	34 (18.8)
Education level			
High school graduate/GED or less	862 (19.7)	50 (10.4)	17 (11.6)
Some college/associate degree	877 (20.1)	77 (16.0)	30 (20.4)
Bachelor’s degree	1,263 (28.9)	167 (34.6)	59 (40.1)
Master’s degree and higher	1,367 (31.3)	188 (39.0)	41 (27.9)

Maternal glycaemic characteristics

- GDM prevalence in the population was 4.6% using the GDM/CC criteria (data not extracted for this population); an additional 3.0% was detected using the GDM/IADPSG for a total of 7.6%
- The mean 50 grams glucose values were significantly different across the four groups and increased across the groups from normal to mild, GDM/IADPSG, and GDM/CC

Study Reference	**Davis 2018**

- Mild hyperglycaemia: Elevated 50 g GCT ≥130 mg/dL and all 4 values normal 3 h 100 g OGTT
- GDM/IADPSG (modified IADPSG): elevated value 1 h 50 g GCT + 1 elevated value on 100 g OGTT using the IADPSG criteria. At this threshold, blood glucose is considered 'normal' under the CC criteria, and women were therefore untreated

Outcomes

Primary endpoint

- LGA birth weight defined as birth weight >90th percentile for GA based on US birth weight standards

Secondary endpoints

- Macrosomia birth weight ≥4000 g
- Primary caesarean delivery defined using current procedural terminology codes 74, 74.1, 74.9
- Hypertensive disorders of pregnancy, a composite variable that included either having pre-eclampsia or gestational hypertension as defined by ICD codes
- Preterm delivery, defined as delivery <37 weeks of gestation
- Severe vaginal lacerations (3rd/4th degree lacerations) • Shoulder dystocia

Glucose tolerance	Normal (n=4,941)	Elevated GCT + NL OGTT (n=544)	GDM by IADPSG criteria (n=181)
50 g GCT value (mg/dL), mean (SD)	99.0 (17.1)	*147.0 (11.9)*	150.6 (13.9)

Methods

Duration of follow-up

NR (retrospective)

Method of blood glucose measurement

Universal screening with a random 50 g GCT for GDM with the majority screened between 24 and 28 weeks' gestation

Diagnostic criteria and test(s)

- 1 h 50 g GCT
- Plus a clinically indicated 3 h 100 g OGTT

Threshold cut-offs

Non-overlapping groups were compared:

- Normal: <130 mg/dL

Study Reference **Davis 2018**

Pregnancy outcomes

Adverse neonatal outcomes

Outcome, n (%)	Normal (n=4,941)	Elevated GCT + Normal OGTT (n=544)					GDM by IADPSG criteria (n=181)				
	Outcome value	Outcome value	Unadjusted OR, CI	pvalue	Adjusted OR, CI	p-value	Outcome value	Unadjusted OR, CI	pvalue	Adjusted OR, CI	p-value
Gestational age at birth, weeks, mean (SD)	39.3 (2.0)	39.3 (2.0)	NR	NR	NR	NR	39.4 (1.9)	NR	NR	NR	NR
Pre-term birth (GA <37 weeks), n (%)	(9.2)	(9.4)	1.020 (0.75 to 1.38)	0.8990	1.243 (0.83 to 1.86)	0.2890	(8.3)	0.891 (0.52 to 1.52)	0.6730	1.423 (0.75 to 2.71)	0.284
Pregnancy complications											
Hypertensive disorder of pregnancy	(8.9)	(9.4)	1.053 (0.78 to 1.43)	0.7396	1.080 (0.70 to 1.66)	0.7227	(12.2)	1.409 (0.89 to 2.22)	0.1406	1.215 (0.63 to 2.35)	0.5627

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)	Normal (n=4,941)	Elevated GCT + NL OGTT (n=544)					GDM by IADPSG criteria (n=181)				
	Outcome value	Outcome value	Unadjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value	Outcome value	Unadjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value
Perinatal mortality	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Mode of birth											
C-section	1,267 (25.6)	175 (32.2)	1.375 (1.14 to 1.66)	0.0011	1.181 (0.91 to 1.52)	0.2023	51 (28.2)	1.138 (0.82 to 1.58)	0.4440	0.810 (0.51 to 1.29)	0.3765
Macrosomia ≥4,000 g	455 (9.2)	59 (10.8)	1.196 (0.90 to 1.59)	0.2217	0.988 (0.66 to 1.48)	0.9551	32 (17.8)	2.126 (1.43 to 3.15)	0.0002	1.876 (1.08 to 3.25)	0.0245
LGA	530 (10.8)	66 (12.1)	1.145 (0.87 to 1.50)	0.3297	0.938 (0.64 to 1.37)	0.7406	34 (18.9)	1.932 (1.32 to 2.84)	0.0008	1.466 (0.85 to 2.53)	0.1708
Birth injury											

Study Reference

Davis 2018

Lacerations (3rd or 4th degree)	(4.2)	(5.5)	1.352 (0.91 to 2.01)	0.1336	1.024 (0.59 to 1.78)	0.9338	(6.7)	1.655 (0.91 to 3.02)	0.1012	0.925 (0.33 to 2.58)	0.8819
Shoulder dystocia	(2.1)	(2.0)	0.953 (0.51 to 1.79)	0.8798	0.540 (0.19 to 1.50)	0.2360	(3.4)	1.592 (0.69 to 3.68)	0.2760	1.294 (0.40 to 4.21)	0.6688
Neonatal hypoglycaemia	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Admission to NICU	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Long-term outcomes	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Study Reference **Delibas 2018**

Study Design	Design Retrospective study
Authors' Conclusions	Women with an elevated 50 g GCT ≥130 mg/dL and a normal 3 h OGTT had similar perinatal outcomes in this study as compared to the women with a normal 50 g GCT (<130mg/dL). The overall low risk of complications among women with a normal 50 g GCT or abnormal 50 g GCT and normal 3 h OGTT suggests that only a very small proportion of women would have had a normal 50 g GCT and significantly elevated 3 h OGTT with downstream adverse outcomes

Abbreviations: BMI, Body Mass Index; Carpenter and Coustan; CI, confidence interval; GA, gestational age; GCT, glucose challenge test; GDM, gestational diabetes mellitus, GED, general educational development; IADPSG, International Association of Diabetes and Pregnancy Study Groups; ICD, International Classification of Disease; LGA, largefor-gestational-age; NL, normal; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; SD, standard deviation; US, United States.

Table 56: Delibas 2018

Study Reference

Delibas 2018

Objective
To determine whether pregnant women who have reactive hypoglycaemia during the 100 g oral glucose tolerance test (OGTT) are at an increased risk of poor pregnancy outcomes.

Dates
January 2012 and December 2014
Country
Turkey
Setting
Obstetric and Clinics Department of Gaziosmanspasa University and Tokat Sate Hospital

Patient recruitment and eligibility

Recruitment: Retrospective review of the perinatal data of all women who underwent a 3h OGTT and gave birth at the study institution during the study period.

Inclusion criteria: Women with singleton pregnancies who had abnormal 1 h 50 g GCT results (≥ 140 mg/dL) at 24–28 weeks of gestation and thus underwent the 3 h 100 g oral GTT were included in the study

Exclusion criteria: The exclusion criteria were twin pregnancies, documented type I or II diabetes mellitus, multiple GCTs in the same pregnancy (only one entry per pregnancy was allowed), and incomplete medical records**. Other:**

Sample size
N screened/invited = NR
N eligible = 421
N enrolled = 413
N excluded (with reason) = 8 (1.9%) due to incomplete medical records

Population Characteristics

N lost to follow-up = NR
N completed = NR
N excluded from analysis = NR
N included in analysis = 413
Maternal demographics

Characteristic	Reactive hypoglycaemia (n=15)	Normo-glycaemia (n=316)	Single high glucose value (n=33)
Age, years, mean (SD)	26.4 (4.4)	28.2 (5.6)	31.4 (5.4)*
Cardiometabolic health			
Pre-pregnant BMI, kg/m^2	NR	NR	NR
BMI, kg/m^2	NR	NR	NR
Weight, kg	NR	NR	NR
Ethnicity, n (%)			
White	NR	NR	NR

Delibas 2018 Study Reference

Black	NR	NR	NR
South Asian	NR	NR	NR
East Asian	NR	NR	NR
Mixed	NR	NR	NR
Medical history/risk factors, n (%)			
Hypertension	NR	NR	NR
Diabetes	NR	NR	NR
Pre-pregnant smoking	NR	NR	NR
Pre-pregnant alcohol use	NR	NR	NR
Obstetric history, mean (SD)			
Parity	0.6 (0.9)	0.6 (0.9)	0.9 (0.9)
Gravida	2.4 (1.1)	2.4 (1.3)	2.7 (1.1)
Education level	NR	NR	NR

Maternal glycaemic characteristics

Glucose tolerance	Study population (N=413)
50 g GCT, mg/dL	
≥140	413 (100)
100 g OGCT, mg/dL	
≤45 mg/dL (reactive hypoglycaemia)	15 (3.6)
All plasma glucose normal (normoglycemia)	316 (76.5)
Single high glucose value, ≥140 mg/dL	33 (0.8)
Gestational diabetes (two or more high glucose values, ≥140 mg/dL)	49 (11.9)

Delibas 2018

Methods	Duration of follow -

NR up

NR lucose measurement

Diagnostic criteria

Non-

nancies are screened for GDM at 24–28 weeks of pregnancy using a two-step standard protocol during a routine prenatal visit. This

- h 50 g GCT, followed by a 3 h 100 g diagnostic OGTT if the GCT plasma glucose result is ≥140 mg/dL
- Reactive hypoglycaemia: glucose ≤45 mg/dL
- Normoglycemia: all plasma glucose values are normal
- Single high glucose value: only one abnormal glucose value

Gestational diabetes: two or more high plasma glucose values

Threshold cut-offs

- GDM was diagnosed when two or more OGTT plasma glucose levels met the criteria for a positive test as recommended by the National Diabetes Data Group (NDDG), which include plasma glucose thresholds of 95 mg/dL for fasting, 180 mg/dL for 1 h, 155 mg/dL for 2h, and 140 mg/dL for 3 h OGTTs

Study Reference

- Reactive hypoglycaemia was defined as a plasma glucose level of < 45 mg/dL (2.5 mmol/L) according to the 1986 Consensus Statement of the Third International Symposium on Hypoglycemia. Another reason for choosing this cut-off plasma glucose level (45 mg/dL) for hypoglycaemia was that it was detected in less than 10% of the study population during OGTTs.

Outcomes

Endpoints (primary and secondary not specified)

- Large-for-gestational-age (LGA) status was defined as a birth weight above the 90th percentile for age
- SGA was defined as a birth weight below the 10th percentile for age
- Macrosomia was defined as an estimated fetal weight of 4,000 g or more, regardless of gestational age
- Apgar score at 5 minutes, weight (g), NICU admission

Pregnancy outcomes

Outcome, n (%)	Reactive hypoglycaemia (n=15)	Normo-glycaemia (n=316)	Single high glucose value (n=33)
Gestational age at birth	37.2 (1.5)	38.5 (1.7)	38.5 (1.3)
Pre-term birth	3 (20.0)	19 (6.0)	3 (9.1)

Authors' Conclusions Although the prevalence of reactive hypoglycaemia during the 3 h 100 g OGTT is relatively low, it is significantly associated with low APGAR scores, low birth weights, and prenatal admission to the NICU

Delibas 2018

Pregnancy complications			
Pre-eclampsia	0 (0)	4 (1.4)	1 (3.3)

Adverse neonatal outcomes

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)	Reactive hypoglycaemia (n=15)	Normo-glycaemia (n=316)	Single high glucose value (n=33)
Perinatal mortality	NR	NR	NR
Mode of birth			
C-section[a]	(28.6)	(28.5)	(33.3)
Macrosomia			
Birth weight, g, mean (SD)	(544.6)	3282.4 (452.8)	3290.6 (510.5)
LGA	(0)	(2.8)	(3.0)
Birth injury	NR	NR	NR
Neonatal hypoglycaemia	NR	NR	NR
Apgar 5 min, mean (SD)	8.3 (1.3)	9.0 (0.8)	8.6 (1.6)
Apgar < 7 (5 min), n (%)	(20.0)	(1.9)	(0)
Admission to NICU	(26.7)	(9.2)	(18.2)
Long-term outcomes	NR	NR	NR

[a] In accordance with guidelines from the Ministry of Health in Turkey, elective C-section was recommended to women with GDM and estimated fetal weights of 4,000 g or more and to women without GDM and estimated fetal weights of 4,500 g or more. Elective and non-elective C-sections were not distinguished between.

Abbreviations: APGAR; appearance, pulse, grimace, activity, and respiration, BMI, Body Mass Index; GCT, glucose challenge test; GDM, gestational diabetes mellitus, GTT, glucose tolerance test; LGA, large-for-gestational-age; NDDG, National Diabetes Data Group; NICU, neonatal/newborn intensive care unit; NR, not reported; OGTT, oral glucose tolerance test; SD, standard deviation.

Table 57: Dennedy 2012 (ATLANTIC-DIP)

Study Reference	Dennedy 2012 (ATLANTIC-DIP)
Study Design	Design Prospective cohort study Objective To investigate the effects of raised maternal BMI on pregnancy outcome in glucose-tolerant women using the International Association of Diabetes and Pregnancy Study Groups criteria Dates September 2006 to 2009 Country Ireland Setting Five antenatal centres
Population Characteristics	Patient recruitment and eligibility **Recruitment:** A representative sample of the obstetric population including euthyroid women with normal glucose tolerance carrying singleton pregnancies was selected **Inclusion criteria:** Euthyroid women with normal glucose tolerance carrying singleton pregnancies were selected **Exclusion criteria:** NR **Other**: NR Sample size N screened/invited = NR N eligible = 3,656 N enrolled = 3,656 N excluded (with reason) = NR N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = 3,656

Maternal demographics

Study Reference

Characteristic	Population (n=3,656)
Age, years, mean (SD; range)	31 (5.3; 16 to 48)

Study Reference **Dennedy 2012 (ATLANTIC-DIP)**

Pre-pregnant BMI, kg/m^2	NR
BMI, kg/m^2, n (%)	
Normal	1,582 (43)
Overweight	1,369 (38)
Obese	695 (19)*
Grade I obese	482 (13)
Grade II obese	168 (5)
Grade III obese	55 (1.5)
Weight, kg	NR
Caucasian, n (%)	3,428 (94)
Hypertension	NR
Diabetes	NR
Pre-pregnant smoking	NR
Smoking during pregnancy, n (%)	291 (8)
Pre-pregnant alcohol use	NR
Nulliparous	NR
Parous without GDM	NR
Parous with GDM	NR
Education level	NR

*Number reported for obese as a whole (695 women) is lower than when different obese classes are added up (705 women).The publication does not refer to this discrepancy.

	Maternal glycaemic characteristics NR
Methods	Duration of follow-up Data was collected from study entry until 12 weeks postpartum Method of blood glucose measurement The 75 g OGTT was performed at 24 to 28 weeks' gestation Diagnostic criteria and test(s) • Fasting plasma glucose • 1 h OGTT • 2 h OGTT Threshold cut-offs Normal glucose tolerance was based on IADPSG recommendations Outcomes Pregnancy outcomes: • Delivery mode (vaginal [normal vs instrumental], lower segment C-section by Pfannen steil incision [LSCS] [elective and emergency])
	Dennedy 2012 (ATLANTIC-DIP)
	• Pregnancy-induced hypertension (PIH; blood pressure >140/90 mmHg on at least two occasions more than 6 h apart in women with normal booking blood pressure) • Preeclamptic toxaemia (PET; hypertension, proteinuria [>300 mg/24 h] onset >20 weeks) • Antepartum haemorrhage (APH; vaginal bleeding from 24 weeks until term) • Postpartum haemorrhage (bleeding >500 mL after vaginal delivery, >1000 mL post-LSCS) Fetal outcomes: • Birthweight • Congenital malformations (ICD-10) • Shoulder dystocia • Neonatal hypoglycaemia • Jaundice • Respiratory distress • Miscarriage (death <20 weeks' gestation) • Stillbirth (death >24 weeks' gestation) • Neonatal death (within 1 week of delivery)

OR per mmol/L of glucose or per SD of glucose

Study Reference	Outcome, OR (95% CI)	Fasting glucose	Glucose 60 min	Glucose 120 min
	Gestational age at birth	NR	NR	NR
	Pre-term birth	NR	NR	NR
	Pregnancy complications			
	Pregnancy-induced hypertension	1.220 (0.663 to 2.246)	1.049 (0.910 to 1.209)	1.160 (0.960 to 1.402)
	Pre-eclamptic toxaemia (hypertension, proteinuria)	0.812 (0.427 to 1.546)	1.157 (0.998 to 1.341)	0.922 (0.759 to 1.120)
	Stillbirth	NR	NR	NR

Adverse neonatal outcomes	Outcome, OR (95% CI)	Fasting glucose	Glucose 60 min	Glucose 120 min
	Perinatal mortality	NR	NR	NR
	Mode of birth			
	Emergency C-section	1.069 (0.672 to 1.699)	1.159 (1.041 to 1.290)[a]	1.094 (0.977 to 1.256)
	Planned C-section	1.400 (0.877 to 2.234)	1.035 (0.930 to 1.152)	0.956 (0.830 to 1.102)
	Macrosomia	1.817 (1.265 to 2.609)[a]	1.065 (0.980 to 1.157)	0.968 (0.867 to 1.082)
	LGA	1.526 (1.034 to 2.253)[b]	1.129 (1.032 to 1.235)[a]	0.954 (0.874 to 1.074)
	Birth injury	NR	NR	NR
	Congenital malformation	0.903 (0.309 to 2.635)	1.095 (0.856 to 3.960)	1.064 (0.770 to 1.472)
	Neonatal hypoglycaemia	NR	NR	NR
	Admission to NICU	NR	NR	NR
	Long-term outcomes	NR	NR	NR

Study Reference	Dennedy 2012 (ATLANTIC-DIP)
	[a]p<0.01; [b]p<0.05
Authors' Conclusions	NR (conclusions related to BMI only)

Abbreviations: BMI: Body Mass Index; CI: confidence interval; GDM: gestational diabetes; IADPSG: Implementation of the International Association of Diabetes and Pregnancy Study Groups; LGA: large-for-gestational age; NICU: neonatal intensive care unit; NR: not reported; OGTT: oral glucose tolerance test; SD: standard deviation

Table 58: Donovan 2017

Study Reference	Donovan 2017
Study Design	Design Retrospective population-based cohort study Objective To examine outcomes associated with alternative glucose thresholds in a 2-step approach for screening and diagnosing GDM Dates October 2008 to December 2012 Country Canada Setting Alberta (universal healthcare system)
Population Characteristics	Patient recruitment and eligibility **Recruitment:** The retrospective cohort study included all pregnancies that occurred during the study period in the province of Alberta, which has approximately 4 million residents **Inclusion criteria:** Pregnant women who had GDM screening, i.e. a 50 g GDM screen followed by a 75 g OGTT when screening was ≥7.8 mmol/L (140 mg/dL) or a 75 g OGTT alone **Exclusion criteria:** Women who delivered prior to 29 weeks of gestation, pregnancies of women with pre-existing diabetes, identified from the APHP antepartum record **Other:** Sample size N screened/invited = 214,254 N eligible = 178,527 N enrolled = 178,527 N excluded (with reason) = 2,217 (delivery prior to 29 weeks' gestation); 162 (birth at unknown gestational age); 1,551 (pre-pregnancy diabetes); 25,969 (no 50 g screen and 75 g OGTT); 5,828 (no subsequent 75 g OGTT) N lost to follow-up = NR N completed = 178,527 N excluded from analysis = NR

Study Reference **Donovan 2017**

N included in analysis = 178,527

Maternal demographics

Characteristic	Normal 50 g screen (N=144,191)	Normal 75 g screen (N=21,248)	HAPO 1.75 (N=4308)
Age, years, mean (SD)	28.8 (5.3)	30.3 (5.3)	31.2 (5.1)
Cardiometabolic health			
Pre-pregnant BMI, kg/m^2	NR	NR	NR
BMI, kg/m^2	NR	NR	NR
Weight, >91 kg, n (%)	12,166 (8.4)	2,077 (9.8)	615 (14.3)
Ethnicity, n (%)			
White	NR	NR	NR
Black	NR	NR	NR
South Asian	NR	NR	NR
East Asian	NR	NR	NR
Mixed	NR	NR	NR
Medical history/risk factors, n (%)			
Hypertension	NR	NR	NR
Diabetes	NR	NR	NR
Smoking	19,611 (13.6)	2,622 (12.3)	529 (12.3)
Pre-pregnant alcohol use	NR	NR	NR
Obstetric history, n (%)			
Nulliparous	64,014 (44.4)	9,195 (43.3)	1,789 (41.5)
Parous without GDM	NR	NR	NR
Parous with GDM	NR	NR	NR

Study Reference **Donovan 2017**

Education level	NR	NR	NR
Median household income, CAD$, mean (SD)	69,305 (19,166)	70,493 (20,109)	70,445.8 (20,568.2)
Urban residence, n (%)	121,228 (84.1)	18,474 (86.9)	3,756 (87.2)

Maternal glycaemic characteristics

Glucose tolerance	**Normal 50 g screen (N=144,191)**	**Normal 75 g screen (N=21,248)**	**HAPO 1.75 (N=4308)**
N (%)	144,191 (80.8)	21,248 (11.9)	4,308 (2.4)
50 g glucose screen, mmol/L	<7.8	≥7.8	≥7.8
75 g OGTT, mmol/L			
Fasting	Not collected	<5.1	≥5.1 to <5.3
1 hour	Not collected	<10.0	≥10 to <10.6
2 hours	Not collected	<8.5	≥8.5 to <9.0

Methods

Duration of follow-up
NR

Study Reference	Donovan 2017
	Method of blood glucose measurement In Alberta, GDM is diagnosed using a 2-step approach, in keeping with the Canadian Diabetes Association 2013 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada A randomly timed 50 g glucose screen is recommended for all pregnant women without previous diagnoses of diabetes by 24 to 28 weeks' gestation, followed by a 75 g OGTT when the screening test is ≥7.8 mmol/L (140 mg/dL) and <10.3 mmol/L (185 mg/dL) Diagnostic criteria and test(s) • 50 g glucose screen • 75 g OGTT Threshold cut-offs • Normal 50 g screen: Normal gestational screens (<7.8 mmol/L). Women were considered to have no GDM and underwent no further testing • Normal 75 g OGTT: Normal 75 g OGTTs, no GDM • HAPO 1.75: at least 1 abnormal value on the 75 g OGTT, corresponding to glucose values associated with an adjusted OR of 1.75 for specified adverse events in the HAPO study and less than an adjusted OR of 2.0. This is the threshold for diagnosis of GDM suggested by IADPSG, albeit without a 50 g screen. Women in this group would not diagnosed with GDM according to the CDA guidelines used in routine practice at the time of study and are therefore unlikely to have been treated Outcomes **Primary endpoint** • LGA rate, defined as having birthweights above the 90th percentile for age and sex on the basis of a national population reference **Secondary endpoints** • Hypertensive disorders of pregnancy, defined as a composite of gestational hypertension, preeclampsia and eclampsia. APHP defines gestational hypertension as a new diastolic blood pressure reading above 90 mm Hg on at least 2 measurements at 20 weeks' gestation, preeclampsia as gestational hypertension with 1+proteinuria or higher on a urinary dipstick recorded on an antepartum risk assessment form, and eclampsia as seizures, as recorded on the intrapartum risk assessment form • Induction of labour • Caesarean delivery • Stillbirth, defined as an infant delivered at 20 weeks of gestation or longer or weighing 500 g or more and without vital signs at birth
	Pregnancy outcomes

Study Reference	Donovan 2017
Adverse neonatal outcomes	

Donovan 2017 Outcome, n (%)	Normal 50 g screen (N=144,191)	Normal 75 g screen (N=21,248)	HAPO 1.75 (N=4,308)
Gestational age at birth	NR	NR	NR
Pre-term birth	NR	NR	NR
Pregnancy complications	NR	NR	NR
Hypertensive disorders of pregnancy	8,028 (5.6)	1,550 (7.3)	(9.1)
Stillbirth	(0.2)	(0.3)	(0.3)

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)	Normal 50 g screen (N=144,191)	Normal 75 g screen (N=21,248)	HAPO 1.75 (N=4,308)
Neonatal death	150 (0.1)	22 (0.1)	6 (0.1)
Mode of birth			
Induction of labour	39,611 (27.5)	5,887 (27.7)	1,274 (29.6)
C-section	37,455 (26.0)	6,535 (30.8)	1,561 (36.2)
Birth weight g, mean (SD)	3,345.6 (538.5)	3,345 (570.6)	3,377.0 (605.7)
Macrosomia (reported as >4000 g)	13,924 (9.5)	2,385 (11.0)	594 (13.5)
LGA	12,045 (8.2)	2,270 (10.5)	628 (14.2)
Birth injury	NR	NR	NR
Neonatal hypoglycaemia	NR	NR	NR
Apgar 5 minutes, 5 to <7, n (%)	3,302 (2.3)	531 (2.4)	122 (2.8)
Admission to NICU	NR	NR	NR
Long-term outcomes	NR	NR	NR

OR of outcomes in each glucose category relative to baseline category

Outcome	Normal 50 g screen (N=144,191)		Normal 75 g screen (N=21,248)		HAPO 1.75 (N=4,308)	
	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
LGA	1 (ref)	NA	1.3 (1.2 to 1.4)	<0.01	1.7 (1.6 to 1.9)	<0.01
C-section	1 (ref)	NA	1.2 (1.1 to 1.2)	<0.01	1.4 (1.3 to 1.5)	<0.01
Induction	1 (ref)	NA	1.0 (1.0 to 1.0)	0.47	1.1 (1.0 to 1.2)	0.01
Hypertensive disorders of pregnancy	1 (ref)	NA	1.3 (1.2 to 1.4)	<0.01	1.5 (1.4 to 1.7)	<0.01

Study Reference	Donovan 2017

Authors' Conclusions	Our results support the use of a 2-step approach for diagnosis of GDM because a negative 50 g screen was associated with a low risk for adverse pregnancy outcomes. However, there was a progressively increased risk for adverse outcomes when the 50 g screen was positive, especially when there was high maternal weight; therefore, the best diagnostic thresholds for the 75 g OGTT remain arbitrary and debatable. Further research is needed to determine whether glycaemic thresholds for GDM diagnosis should incorporate information on maternal weight

Abbreviations: CI: confidence interval; GDM: gestational diabetes; LGA: large-for-gestational age; NICU: neonatal intensive care unit; OR: odds ratio; OGTT: oral glucose tolerance test; OR: odds ratio; SD: standard deviation.

Table 59: Ezell 2015

Study Reference	Ezell 2015
Study Design	Design Prospective cohort study

Study Reference	**Ezell 2015**
	Objective To examine the association between 1 h glucose challenge test (GCT) values and risk of caesarean section Dates February 2009 to June 2010 Country US Setting Obstetric clinics in the Henry Ford Health System in metropolitan Detroit, Michigan
	Patient recruitment and eligibility **Recruitment:** NR **Inclusion criteria:** Pregnant black women between the ages of 18 to 44, receiving prenatal care from obstetric clinics **Exclusion criteria:** NR **Other:** NR Sample size N screened/invited = 203 N eligible = 158 N enrolled = 158 N excluded (with reason) = clinician-documented GDM in the current pregnancy (n=12); 3 h 100 g OGTT values consistent with GDM (n=7); preexisting type 2 diabetes (n=5); never screened for GDM (n=5); incomplete 1 h GCT (n=2); abnormally high 1 h GCT result never followed up for diagnostics, due to presentation of labour (n=2); met Leykin and Pellis (2009) definition for "super-super" morbid obesity (maternal pre-pregnancy BMI >60 kg/m^2; n=3); twin pregnancy (n=2); missing information due to delivery occurring at an outside facility (n=7)
Population Characteristics	N lost to follow-up = NR N completed = 158 N excluded from analysis = 0 N included in analysis = 158 Maternal demographics

Study Reference **Ezell 2015**

Characteristic	Vaginal delivery (n=105)	C-section delivery (n=53)	P value
	25.9 (6.1)	26.1 (5.6)	0.825
Pre-pregnant BMI, kg/m , mean (SD)	27.2 (6.7)	31.4 (7.3)	<0.001
Underweight, n (%)	6 (5.7)	1 (1.9)	NR
Normal weight, n (%)	37 (35.2)	12 (22.6)	NR
Overweight, n (%)	35 (33.3)	8 (15.1)	NR
Obese, n (%)	27 (25.7)	32 (60.4)	NR
BMI, kg/m²	NR	NR	NR
Weight, kg	NR	NR	NR

Age, years, mean (SD) Cardiometabolic health

Ethnicity, n (%)	All women were black as per the inclusion criteria		
Medical history/risk factors, n (%)			
Hypertension	NR	NR	NR
Diabetes	NR	NR	NR
Pre-pregnant smoking	NR	NR	NR
Smoking during pregnancy	9 (8.6)	4 (7.5)	0.545
Pre-pregnant alcohol use	NR	NR	NR
Obstetric history, n (%)			
Nulliparous	47 (44.8)	29 (54.7)	0.237
Parous without GDM	NR	NR	NR
Parous with GDM	NR	NR	NR
Education level, years, mean (SD)	12.8 (1.8)	12.9 (1.4)	0.888

Maternal glycaemic characteristics

The overall mean 1 h GCT value was 104.2 (SD 21.3) mg/dL

Study Reference	**Ezell 2015**
Methods	Duration of follow-up NR Method of blood glucose measurement In accordance with the ACOG guidelines, as part of routine prenatal care, women were screened for GDM at approximately 28 weeks of gestation using the 1 h 50 g GCT. Women classed as screening "positive" were then tested for GDM with the 3 h 100 g OGTT Diagnostic criteria and test(s) • 1 h 50 g GCT: The lower bound for screening positive depended on the individual medical provider making the determination; at HFHS, the criteria to classify women as screening positive varied slightly (cut-offs of GCT ≥130 mg/dL, ≥135 mg/dL, or ≥140 mg/dL were used by different clinicians. For purposes of analysis, the primary analysis was done using continuous GCT levels; when examining based on categorical considerations, an abnormal GCT screen was defined using the mid-point of the value used at HFHS of ≥135 mg/dL • 3 h 100 g OGTT: Unclear, but appears GDM was diagnosed using ACOG criteria Threshold cut-offs • OR for 1 mg/dL increments in glucose level measured by 1 h 50 g GCT • Elevated glucose level by 1 h 50 g GCT: >135 mg/dL Logistic regression models were fit to examine the association of continuously distributed 1 h GCT values and delivery mode (vaginal versus C-section). Models were fit unadjusted and then adjusted for potential confounding variables, specifically maternal age, previous C-section, and maternal prepregnancy BMI, which were identified in the literature as variables associated with delivery mode and/or 1 h GCT value. The authors then refit their models stratified by parity status (nulliparous compared to parous) Outcomes **Primary endpoint** • C-section **Secondary endpoints** None reported in relation to 50 g GCT glucose levels
	Pregnancy outcomes NR Neonatal outcomes according to maternal glycaemic status

Ezell 2015
NR

Study Reference

OR of outcomes in each glucose category relative to baseline category

- In contrast to women with 1 h GCT values <135 mg/dL, parous women with an elevated 1 h GCT were at 5.1 times higher odds of having a C-section (95% CI: 0.7 to 37.4; p = 0.113), after adjusting for maternal age, maternal pre-pregnancy BMI and prior C-section

OR per 1 mg/dL increments in glucose 50 g GCT value and delivery by C-section, compared to vaginal delivery, in the overall sample and stratified by parity

Outcome, n (%)	Overall		Nulliparous		Parous	
	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
h GCT (mg/dL): Model 1	1.01 (1.00 to 1.03)	0.131	1.00 (0.98 to 1.02)	0.856	1.03 (1.00 to 1.05)	0.034
h GCT (mg/dL): Model 2	1.01 (1.00 to 1.03)	0.171	1.00 (0.98 to 1.02)	0.943	1.05 (1.01 to 1.09)	0.017
h GCT (mg/dL): Model 3	1.01 (0.99 to 1.03)	0.356	1.00 (0.98 to 1.02)	0.884	1.05 (1.00 to 1.05)	0.029

Adverse neonatal outcomes

In multivariable models, there was no evidence of an association between unplanned C-section and 1 h GCT values (relative to vaginal delivery). Conversely, for every 1 mg/dL increase in 1 h GCT value, the unadjusted odds of having a planned C-section versus vaginal delivery increased by 1.03 (95% CI: 1.00 to 1.05; p = 0.036). Adjusting for maternal age, previous C-section, and maternal pre-pregnancy BMI, the association between 1 h GCT value and risk of planned C-section compared to vaginal delivery was borderline statistically significant (OR = 1.06; 95% CI: 1.0 to 1.12; p = 0.051) (models not shown in publication)

	Ezell 2015
Authors' Conclusions	In multivariable models, there was no evidence of an association between unplanned C-section and 1 h GCT values (relative to vaginal delivery). Conversely, for every 1 mg/dL increase in 1 h GCT value, the unadjusted odds of having a planned C-section versus vaginal delivery increased by 1.03 (95% CI: 1.00 to 1.05; p = 0.036). Adjusting for maternal age, previous C-section, and maternal pre-pregnancy BMI, the association between 1 h GCT value and risk of planned C-section compared to vaginal delivery was borderline statistically significant (OR = 1.06; 95% CI: 1.0 to 1.12; p = 0.051) (models not shown in publication)

Abbreviations: ACOG; American College of Obstetricians and Gynecologists; BMI, Body Mass Index; CI, confidence interval; GCT, glucose challenge test; GDM, gestational diabetes mellitus; HFHS, Henry Ford Health System; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; SD, standard deviation; US, United States.

Table 60: Jiang 2017

Study Reference	Jiang 2017
Study Design	Design Retrospective cohort study Objective To assess the pregnancy-related outcomes of women according to the different diagnostic criteria for GDM adjusting for body mass index categories. Dates January 2011 and April 2015 Country Australia Setting Westmead Hospital Institute Clinical Pathology and Medical Research database
Population Characteristics	Patient recruitment and eligibility **Recruitment:** All women satisfying the inclusion criteria of singleton pregnancy, having an antenatal 75 g OGTT after 20 weeks of pregnancy, no history of pre-gestational diabetes and delivery >24 weeks' gestation at the study institution during the study period were included **Other:** Only the first pregnancy was used for analysis if a woman had multiple pregnancies during this period Sample size N screened/invited = NR N eligible = 4081 N enrolled = NR N excluded (with reason) = NR N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = NR Maternal demographics

Characteristic	Control (N=3185)	GDM 2010-Only (untreated; N=94)

Age, years, median (IQR)	30.0 (27 to 33)	30.5 (27 to 34)
Maternal body BMI, *kg/m^{2,}* median (IQR)	23.44 (20.27 to 26.62)	25.78 (21.71 to 29.85)
Pre-pregnant BMI, kg/m^2	NR	NR
Weight by category		
Normal	1747 (55.2)	41 (44.1)
Underweight	226 (7.1)	2 (2.2)
Overweight	711 (22.5)	23 (24.7)
Obese	480 (15.2)	27 (29.0)
Caucasian	1435 (45.1)	35 (37.2)

Study Reference **Jiang 2017**

Subcontinental	882 (27.7)	42 (44.7)
East and Southeast Asian	590 (18.5)	8 (8.5)
African	112 (3.5)	4 (4.3)
South American	38 (1.2)	1 (1.1)
Polynesian	126 (4.0)	4 (4.3)
Medical history, n (%)		
Hypertension	NR	NR
Family history of diabetes	1169 (36.9)	43 (45.7)
Smoking in pregnancy	186 (5.8)	3 (3.2)
Pre-pregnant alcohol use	NR	NR
Nulliparous	NR	NR
Parous without GDM	NR	NR
Parous with GDM	NR	NR
Education level	NR	NR

Maternal glycaemic characteristics

Glucose tolerance[a]	Control (N=3185)	GDM 2010-Only (N=94)

FPG, mmol/L	4.20 (3.25 to 5.15)	5.2 (5.1 to 5.3)
2-hour 75 g OGTT, mmol/L	6.1 (5.35 to 6.85)	6.6 (5.95 to 7.25)

[a] Kruskal-Wallis

Methods	Duration of follow-up N/A – data abstracted from database of obstetric outcomes Method of blood glucose measurement During this period, standard antenatal care at Westmead Hospital for all women reaching 24–28 weeks gestation included a one-hour non-fasting 50 g GCT. Those with a one-hour result ≥7.8 mmol/L would be referred to have a two sample 75 g OGTT. This involves measuring venous plasma BGL from peripheral venous blood sampling at fasting, and two hours post-ingestion of a 75 g glucose load. Diagnostic criteria and test(s) • Fasting blood glucose • 75 g OGTT The diagnosis of GDM and referral for subsequent management was based on the ADIPS 1998 criteria (fasting glucose ≥55 mmol/L and/or two hours' result ≥8.0 mmol/L). Threshold cut-offs • Control cohort: women without GDM on any diagnostic criteria on 75 g OGTT (fasting BGL <5.1 and 2 hours BGL <8.0) • 'GDM 2010-Only' group: women who would be diagnosed with GDM according to the new IADPSG 2010 criteria only but did not satisfy the ADIPS 1998 criteria (fasting BGL 5.1 to 5.4 mmol/L and two hours BGL <8.0 mmol/L) Outcomes • LGA, defined as >90th neonatal birth centile

Outcome, n (%)	Control	GDM 2010-Only	OR (95% CI)
Gestational age at birth	NR	NR	NR
Pre-term birth	(5.6)	(4.3)	0.75 (0.27 to 2.07)
Pregnancy complications			
Pre-eclampsia	NR	NR	NR

Study Reference	*Stillbirth*	(0.3)	(0)	-	**Jiang 2017**

- SGA, defined as <10th neonatal birth centile
  - Neonatal birth centile was calculated using a customised birthweight centile calculator adjusting for maternal age, parity, ethnicity, neonatal sex, gestation and weight
- Preterm birth, defined as delivery <37 weeks gestation
- Primary C-section was only analysed in women without a previous C-section or major uterine surgery in order to avoid confounders such as the need for a C-section due to a previous uterine scar
- Shoulder dystocia, the application of McRobert's manoeuvre was used as an indicator of the presence of shoulder dystocia

Adverse neonatal outcomes

Pregnancy outcomes

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)	Control	GDM 2010-Only	OR (95% CI)
Perinatal mortality	NR	NR	NR
Mode of birth	NR	NR	NR
Primary C-section	536 (20.1)	24 (33.8)	2.03 (1.23 to 3.35)
Macrosomia	NR	NR	NR
LGA	298 (9.4)	19 (20.2)	2.45 (1.46 to 4.12)
Birth injury			
Shoulder dystocia	215 (6.8)	*5 (5.3)*	0.78 (0.31 to 1.93)
Brachial plexus neuropathy	NR	NR	NR
Neonatal hypoglycaemia	NR	NR	NR
Admission to NICU	NR	NR	NR
Long-term outcomes	NR	NR	NR

Authors' Conclusions Untreated women who would be diagnosed with GDM using the new criteria have an increased risk of pregnancy complications, with maternal obesity having an even greater risk.

Abbreviations: ADIPS, Australasian Diabetes in Pregnancy Society**;** BGL, blood glucose level; BMI, body mass index; CI, confidence interval; FPG, fasting plasma glucose; GDM, gestational diabetes; IQR, interquartile range; LGA, large-for-gestational age; NICU, neonatal intensive care unit; N/A, not applicable; NR, not reported; OGTT, oral glucose tolerance test; SGA, small-for-gestational age

Table 61: Lopez de Val 2019 (GDMFU)

Study Reference	Lopez de Val 2019
Study Design	Design Retrospective study Objective To establish whether fasting glucose levels in the first trimester (FGFT) of pregnancy ≥92 mg/dl (5.1 mmol/l) (FGFT) anticipate the occurrence of maternal-fetal complications of gestational diabetes mellitus. To assess whether FGFT can replace diagnosis of GDM using the classical two-step oral glucose tolerance test (OGTT) Dates NR Country Spain Setting Obstetrics Department of Hospital Severo Ochoea, Madrid
	Patient recruitment and eligibility **Recruitment:** The GDMFU study was an unplanned follow-up study of women, and their offspring who participated in the Eunice Kennedy Shriver National Institute of Child Health and Human Development's Maternal-Fetal Medicine Units (MFMU)

Exclusion criteria: Patients with prior diabetes or who had prenatal screening but suffered subsequent miscarriage

Other: NR Sample size

N screened/invited = 1425
N eligible = 1425
N enrolled = NR
N excluded (with reason) = 59 (miscarriage following screening)

	Weight, kg		(22.3)	NR
	Ethnicity, n (%)	(11.9)		NR
	Medical history/risk factors, n (%)		NR	
	Pre-gestational hypertension	NR	NR	0.09
	Smokers	NR		0.2
	Family history of diabetes		(3.1) (11.9)	<0.01
	Previous GDM	(1.4) (15.1)	(29.0)	<0.01
	Previous miscarriages	(16.7) (3.0	(8.6)	0.5
	Pre-pregnant alcohol use	(28.3) NR	(27.9) NR	NR
	Obstetric history, n (%)			
	Previous pregnancies	(40.4)	(53.8)	<0.01
	Parous without GDM	NR	NR	NR
	Parous with GDM	NR	NR	NR
	Education level	NR	NR	NR

Population Characteristics

N lost to follow-up = NR
N completed = NR
N excluded from analysis = 84 women diagnosed with GDM and therefore treated for GDM N included in analysis = 1348

Maternal demographics

Inclusion criteria: The population reported in this analysis consisted of offspring of untreated mild GDM as well as offspring of non-GDM women.

Age, years, mean (SD)	32.4 (5.1)	33.8 (4.4)	<0.01
Cardiometabolic health			
Pre-pregnant BMI, kg/m^2	NR	NR	NR
BMI, kg/m^2, mean (SD)	24.6 (4.5)	26.6 (5.0)	<0.01
<25, n (%)	659 (64.2)		<0.01
25–29.9, n (%)	246 (24.0)	63 (40.1)	

Characteristic	FGFT <92 (n=1239)	FGFT ≥92 (N=193)	p-value
Study Reference	**Lopez de Val 2019**		
≥30, n (%)			

	Maternal glycaemic characteristics NR
Methods	Duration of follow-up NR Method of blood glucose measurement First trimester fasting blood glucose (FGFT) and O'Sullivan test (OST) in weeks 24 to 28 of gestation. In the OST-positive patients the OGTT 100 g results were also compiled. Diagnostic criteria and test(s) The Coustan-Carpenter criteria was used for the diagnosis GDM (these patients were therefore treated but are not included in the outcome analysis) Threshold cut-offs Outcomes were related to the glycaemia groups of <92 mg/dL and ≥92 mg/dL. Outcomes **Obstetric endpoints** • Gestational hypertension • Pre-eclampsia • Polyhydramnios • Pre-term delivery, defined as delivery before week 37 • Delivery to term (spontaneous, induced, eutocic, instrumental) • C-section **Neonatal endpoints** • Macrosomia, defined as weight >4000 g • Hyperbilirubinaemia • Polyglobulia • Hypoglycaemia

Study Reference

Lopez de Val 2019

- Intrauterine fetal death
- Malformations
- Trauma to the newborn resulting from labour
- Respiratory distress
- Admission to the NICU

Pregnancy outcome

Results referring to obstetric and fetal complications according to fasting glucose in the first trimester (FGFT) ≥92 and <92 mg/dl in pregnant women not diagnosed with diabetes

Gestational age at birth	NR	NR	NR	
Pre-term birth	NR	NR	NR	
Pregnancy complications *Pre-eclampsia*	19 (1.7)	3 (2.1)	1.02 (0.28 to 3.75, pvalue ns	
Stillbirth	NR	NR	NR	

Neonatal outcomes according to c status maternal glycaemi

Outcome, n (%)		FGFT <92 (n=1193)	FGFT ≥92 (N=155)	OR (95% CI), pvalue	AOR (95% CI), pvalue
	Outcome, n (%)	FGFT <92 (n=1193)	FGFT ≥92 (N=155)	OR (95% CI), pvalue	
Adverse neonatal outcomes	**Perinatal mortality**	NR	NR	NR	NR
	Mode of birth	NR	NR	NR	NR
	Macrosomia, n (%)	40 (3.4)	12 (7.2)	2.42 (1.27 to 4.62), p<0.05	1.50 (0.63 to 3.57)[a] p=0.3
	LGA	NR	NR	NR	NR
	Birth injury				

Trauma during vaginal delivery	19 (1.6)	9 (5.7)	3.10 (1.15 to 8.32), p=0.02	NR
Brachial plexus neuropathy	NR	NR	NR	NR
Respiratory distress	54 (4.6)	8 (5.1)	1.03 (0.34 to 3.21)	NR
Neonatal hypoglycaemia	26 (2.2)	4 (2.6)	0.98 (0.27 to 3.55), NR	pvalue ns
Admission to NICU, n (%)	134 (11.2)	25 (16.5)	1.60 (0.94 to 2.89),[b] p=0.08	1.50 (0.78 to 2.73), pvalue ns
Long-term outcomes	NR	NR	NR	NR

[a] Adjusted for BMI, age and previous GDM [b] Adjusted for BMI and C-section

Authors' Conclusions	Pregnant women with FGFT levels ≥92 mg/dl, even with no subsequent diagnosis of GDM, are a risk group for fetal macrosomia and could benefit from dietary measures and physical exercise.

Abbreviations: aOR: adjusted odds ratio; CI: confidence interval**;** FGFT: fasting glucose first trimester; GDM: gestational diabetes; NICU: neonatal intensive care unit; NR: not reported; NS: non-significant; OR: odds ratio; SD: standard deviation.

Table 62: Meek 2015

Study Reference	**Meek 2015**
Study Reference	**Meek 2015**
Study Design	Design Retrospective study Objective To assess neonatal and obstetric outcomes among women who test positive for the IADPSG criteria but negative for the NICE 2015 criteria Dates 2004 to 2008 Country England, UK Setting Cambridge University Hospitals National Health Service Foundation Trust

Patient recruitment and eligibility

Recruitment: All pregnant women during the study period were invited to be screened at antenatal booking.
Inclusion criteria: Further eligibility criteria were not reported
Exclusion criteria: Women with pre-existing diabetes

Other: NR

Sample size

N screened/invited = 25,789
N eligible = 25,543
N enrolled = NR
N excluded (with reason) = 246; miscarriage (n=59), termination (n=65), no birthweight information (n=3), duplicate data (n=20), records consistent with overt diabetes (RPG ≥11.1 mmol/l at booking (n=99) N lost to follow-up = NR

tics

N completed = NR
N excluded from analysis = 0 N included in analysis =25,543

Maternal demographics

Characteristic	NICE-negative IADPSG-negative (N=2406)	NICE-negative IAD-positive (N=387)	IADPSG-only 0 hr (N=167)	IADPSG-only 1 hr (N=288)
Maternal age, years, mean (95% CI)	31.4 (31.2 to 31.6)***	32.6 (32.1 to 33.1)***	32.7 (31.9 to 33.5)***	32.6 (32.0 to 33.2)***
Cardiometabolic health				
Maternal usual BMI, kg/m^2, mean (95% CI)	26.0 (25.7 to 26.2)***	27.4 (26.8 to 28.1)***	29.0 (27.9 to 30.1)***	27.2 (26.5 to 27.9)***
Weight, kg	NR	NR	NR	NR
Ethnicity, n (%)				
White	2,151 (89.4)	336 (86.8)	144 (86.2)	246 (85.4)
Black	28 (1.2)	4 (1.0)	1 (0.6)	3 (1.0)
Asian	126 (5.2)	36 (9.3)	18 (10.8)	31 (10.8)
Other	89 (3.7)	11 (2.8)	4 (2.4)	8 (2.8)
Medical history/risk factors, n (%)				
Hypertension	NR	NR	NR	NR
Diabetes	NR	NR	NR	NR
Maternal smoking	177 (7.4)	27 (7.0)	11 (6.6)	21 (7.3)
Pre-pregnant alcohol use				
Obstetric history, n (%)				
Primiparous	941 (39.1)	141 (36.4)	57 (34.1)	103 (35.8)
Nulliparous	NR	NR	NR	NR
Education level	NR	NR	NR	NR

*** p<0.001 by linear regression

Maternal glycaemic characteristics NR

Study Reference	Meek 2015
Methods	Duration of follow-up Until delivery Method of blood glucose measurement All pregnant women were invited to be screened at antenatal booking with a random plasma glucose, typically at 12–16 weeks' gestation. Both venous and capillary samples were used during 2004 and 2008 for glucose testing. Diagnostic criteria and test(s) All women were screened with a random plasma glucose (RPG) test. Women with RPG>7.0 mmol/L or a previous GDM diagnosis were offered a 75 g OGTT. All women without known GDM/pre-existing diabetes were screened at 26–28 weeks with a 50 g glucose challenge test (GCT). Women with a GCT result >7.7 mmol/l were then referred for a 75 g OGTT. Therefore, all women who had an OGTT (n=3,848) had already had at least one abnormal result on glucose testing during pregnancy, symptoms consistent with hyperglycaemia, or GDM in a previous pregnancy. The WHO 1999 criteria were used for GDM diagnosis until August 2007 (75 g OGTT 0 h ≥7.1 mmol/L; 2 h ≥7.8 mmol/L) and the modified WHO 1999 criteria thereafter (75 g OGTT 0 h ≥6.1 mmol/L; 2 h ≥7.8 mmol/l). The criteria proposed by NICE were 75 g OGTT 0 h ≥5.6 mmol/l; 2 h ≥7.8 mmol/L. Threshold cut-offs **IADPSG criteria**: 75 g OGTT 0 h ≥5.1 mmol/l, 1 h ≥10.0 mmol/l, 2 h ≥8.5 mmol/l **NICE 2015 criteria**: 75 g OGTT 0 h ≥5.6 mmol/l; 2 h ≥7.8 mmol/l Groups • **NICE-negative IADPSG-negative (N=2,406)**: negative for GDM by both criteria • **NICE-2015 negative, IADPSG-positive (N=387)**: IADPSG-only 0 h OGTT value 5.1–5.5 mmol/l. Treatment offered to 0% • **IADPSG-only 0 hr (N=167)**: ≥5.1 mmol/l on 75 g 1-hr OGTT. Treatment offered to 0% • **IADPSG-only 1 hr (N=288)**: ≥10.0 mmol/l on 75 g 1-hr OGTT. Treatment offered to 0%

Outcomes

- Macrosomia, defined as birthweight >4000 g
- LGA, defined as birthweight >90th percentile for gestational age and was calculated for babies at 24 to 41 weeks' gestation using the WHO weight percentile calculator with a mean birthweight of 3,542 g (SD 437 g)
- Pre-eclampsia, defined as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg on two or more occasions with proteinuria ≥1+ on dipstick. Women with chronic hypertension prior to pregnancy were not considered to have pre-eclampsia
- Preterm delivery was defined as delivery prior to 37 week's gestation
- Antepartum haemorrhage, defined as any blood loss form the vagina after the 24th week of gestation
- Postpartum haemorrhage, defined as blood loss of >500 ml following delivery, or the requirement for a blood transfusion

Pregnancy outcomes

Study Reference
Adverse neonatal outcomes

Meek 2015 Outcome, n (%)	NICE-negative IADPSGnegative (N=2,406)			NICE-negative IAD-positive (N=387)			IADPSG-only 0 hr (N=167)			IADPSG-only 1 hr (N=288)		
	n (%) or mean (95%CI)	OR (95% CI)	AOR (95% CI)	n (%) or mean (95%CI)	OR (95% CI)	AOR (95% CI)	n (%) or mean (95%CI)	OR (95% CI)	AOR (95% CI)	n (%) or mean (95%CI)	OR (95% CI)	AOR (95% CI)
Gestational age at birth	39.3 (39.3 to 39.4)	NA	NA	39.1 (38.9 to 39.2)	NA	NA	39.1 (38.8 to 39.4)	NA	NA	39.1 (38.8 to 39.3)	NA	NA
Pre-term birth	127 (5.3)	0.75 (0.63 to 0.91)	0.72 (0.59 to 0.89)	(7.5)	1.10 (0.75 to 1.61)	1.02 (0.68 to 1.55)	(6.6)	0.95 (0.52 to 1.76)	0.88 (0.46 to 1.71)	(7.3)	1.06 (0.68 to 1.67)	0.97 (0.60 to 1.57)
Pregnancy complications												
Pre-eclampsia	174 (7.2)	1.40 (1.18 to 1.65)	1.21 (1.01 to 1.44)	(10.1)	2.01 (1.43 to 2.81)	1.40 (0.97 to 2.03)	(9.6)	1.90 (1.13 to 3.19)	1.12 (0.63 to 1.99)	(11.1)	2.24 (1.53 to 3.25)	1.66 (1.11 to 2.48)
Stillbirth	(0.2)	0.58 (0.24 to 1.45)	1.09 (0.42 to 2.79)	(0.3)	0.73 (0.10 to 5.24)	1.16 (0.16 to 8.70)	(0)	Insufficient events	Insufficient events	(0.3)	0.98 (0.14 to 7.06)	1.46 (0.19 to 11.09)

OR adjusted for maternal BMI, maternal age, parity, maternal smoking, ethnicity (pre-eclampsia) + pre-eclampsia, antepartum haemorrhage (stillbirth) Reference for ORs is OGTT not done population

Neonatal outcomes according to maternal glycaemic status

NICE-negative IADPSG- (N=167)	NICE-negative IAD-positive (N=288)	IADPSG-only 0 hr	IADPSG-only 1 hr	negative (N=2406)

	n (%)	OR (95% CI)	AOR (95% CI)	n (%)	OR (95% CI)	AOR (95% CI)	n (%)	OR (95% CI)	AOR (95% CI)	n (%)	OR (95% CI)	AOR (95% CI)
Perinatal mortality	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Mode of birth												
Induction of labour	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Study Reference **Meek 2015**

Spontaneous vertex delivery (vaginal)	1,264 (52.5)	NR	NR	190 (49.1)	NR	NR	74 (44.3)	NR	NR	147 (51.0)	NR	NR
Instrumental delivery	324 (13.5)	1.05 (0.93 to 1.19)	1.06 (0.93 to 1.21)	50 (12.9)	1.00 (0.74 to 1.35)	1.00 (0.72 to 1.40)	20 (12.0)	0.92 (0.57 to 1.47)	0.94 (0.55 to 1.61)	38 (13.2)	1.02 (0.73 to 1.44)	1.03 (0.70 to 1.52)
Emergency Csection	473 (19.7)	1.45 (1.30 to 1.61)	**1.31 (1.16 to 1.47)**	94 (24.3)	1.90 (1.50 to 2.41)	**1.60 (1.24 to 2.06)**	44 (26.3)	2.12 (1.50 to 3.00	**1.66 (1.13 to 2.43)**	68 (23.6)	1.81 (1.39 to 2.41)	**1.49 (1.10 to 2.01)**
Planned C-section	342 (14.2)	NR	NR	53 (13.7)	NR	NR	29 (17.4)	NR	NR	35 (12.2)	NR	NR
Other/unknown	3 (0.1)	NR	NR	0 (0)	NR	NR	0 (0)	NR	NR	0 (0)	NR	NR
Macrosomia	403 (16.8)	1.60 (1.42 to 1.79)	**1.52 (1.34 to 1.73)**	112 (28.9)	3.23 (2.59 to 4.04)	**3.55 (2.75 to 4.58)**	61 (36.5)	4.57 (3.33 to 6.28)	**5.02 (3.46 to 7.28)**	77 (26.7)	2.90 (2.22 to 3.77)	**3.21 (2.38 to 4.34)**
LGA	406 (16.9)	1.75 (1.56 to 1.96)	**1.63 (1.44 to 1.84)**	115 (29.7)	3.64 (2.91 to 4.56)	3.12 (2.44 to 3.98)	63 (37.7)	5.24 (3.81 to 7.21)	**4.47 (3.15 to 6.33)**	75 (26.0)	3.04 (2.33 to 3.98)	**2.58 (1.93 to 3.46)**
Birth injury	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Neonatal hypoglycaemia	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
1 min Apgar score <7, n (%)	141 (5.9)	0.98 (0.82 to 1.17)	1.05 (0.86 to 1.28)	34 (8.8)	1.51 (1.06 to 2.16)	**1.55 (1.06 to 2.26)**	18 (10.8)	1.88 (1.15 to 3.08)	**2.16 (1.30 to 3.60**	25 (8.7)	1.49 (0.99 to 2.26)	1.48 (0.94 to 2.31)
5 min Apgar score <7, n (%)	11 (0.5)	0.53 (0.29 to 0.97)	0.76 (0.40 to 1.47)	4 (1.0)	1.21 (0.45 to 3.27)	1.36 (0.43 to 4.38)	1 (0.6)	0.69 (0.10 to 4.98)	1.03 (0.14 to 7.53)	4 (1.4)	1.63 (0.60 to 4.42)	1.88 (0.58 to 6.05)

Authors' The IADPSG criteria identify women at substantial risk of complications who would not be identified by the NICE 2015 criteria. **Conclusions**

Abbreviations: AOR, adjusted odds ratio; BMI, body mass index; CI, confidence interval; GCT, glucose challenge test; GDM, gestational diabetes; IADPSG, International

	Study Reference	Meek 2015											
	Admission to NICU	143 (5.9)	0.86 (0.72 to 1.02)	1.09 (0.89 to 1.35)	22 (5.7)	0.82 (0.53 to 1.26)	0.76 (0.45 to 1.29)	10 (6.0)	0.86 (0.45 to 1.64)	1.05 (0.51 to 2.14)	16 (5.6)	0.80 (0.48 to 1.33)	0.65 (0.34 to 1.23)
	Long-term outcomes	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Reference for ORs is OGTT not done population

Association of Diabetes and Pregnancy Study Groups; LGA, large-for-gestational age; NICU, neonatal intensive care unit; NA, not available; NICE, The National Institute for Health and Care Excellence; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; RPG, random plasma glucose; SD, standard deviation; UK, United Kingdom; WHO, World Health Organization.

Table 63: Miyakoshi 2010

Study Reference	**Miyakoshi 2010**
Study Design	Design Retrospective cohort study Objective To investigate the metabolic phenotype and pregnancy outcomes of gestational impaired glucose tolerance (IGT) defined by isolated hyperglycaemia during an OGTT Dates January 1996 to August 2008 Country Japan Setting Keio University (referral) hospital
Population Characteristics	Patient recruitment and eligibility **Recruitment:** It appears that all women from the authors' institution fulfilling the inclusion criteria were retrospectively included. **Inclusion criteria:** Women who underwent universal screening for GDM at the authors' hospital **Exclusion criteria:** Women with multiple pregnancies, pregnancies with congenital abnormalities, history of glucose intolerance or of the use of medications known to affect glucose metabolism. **Other:** It is unclear where the data on glucose levels was abstracted from, but data on pregnancy and maternal outcomes was taken form women's hospital records. Sample size

Study Reference	N screened/invited = NR

Miyakoshi 2010

N eligible = 4,789
N enrolled = 1,025

N excluded (with reason) = NR

N lost to follow-up = NR
N completed = 1,025 N excluded from analysis = NA N included in analysis =1,025

Maternal demographics

Reported for women screened 1996 to 2008

Characteristic	Normal OGTT (N=200)	2 h IGT (N=26)	1 h IGT (N=18)
Age, years, mean (SD)	33.2 (4.5)	34.5 (3.8)**	35.2 (3.8)**
Cardiometabolic health			
Pre-pregnant BMI, kg/m2, mean (SD)	20.3 (2.4)	20.9 (3.2)**	20.6 (2.7)
Overweight (BMI ≥25 kg/m^2) (%)	5.3	10.2*	6.1
Underweight (BMI ≥25 kg/m^2) (%)	21.4	13.9	19.7
Ethnicity, n (%)	NR	NR	NR
Medical history/risk factors, n (%)			
Diabetes (%)	6.1	14.8**	12.1
Obstetric history, %			
Parous	30.1	34.2	31.8
Parous with GDM	0.4	0.9	3.1*
Education level	NR	NR	NR

*p<0.05 vs normal glucose, **p<0.01 vs normal glucose

Methods

Duration of follow-up
Until birth

Method of blood glucose measurement
1 h 50 g GCT: venous blood sample 1 hour after ingesting 50 g of glucose, administered regardless of fasting
2 h 75 g OGTT: venous blood sample in fasting state (12 h overnight fast), 30 min, 1 h and 2 h after ingestion of 75 g glucose. Diagnostic criteria and test(s)

- 1 h 50 g GCT to identify women to receive the OGTT
- GDM and IGT diagnosed based on Japan Society of Obstetrics and Gynaecology criteria using 2 h 75 g OGTT

Threshold cut-offs

- GCT >7.8 mmol/L
- IGT one value above: 5.6 mmol/L fasting, 10.0 mmol/L 1 h, 8.3 mmol/L 2 h
- GDM ≥2 abnormal values: 5.6 mmol/L fasting, 10.0 mmol/L 1 h, 8.3 mmol/L 2 h

Study Reference **Miyakoshi 2010**

Outcomes
Primary outcome is implied to be LGA, with other endpoints being gestational hypertension, pre-eclampsia, macrosomia. SGA was not pre-specified but also reported.

Pregnancy outcomes

Reported for 4,789 women screened 1996 to 2008

Outcome	Normal glucose tolerance (N=4,512)	2 h IGT (N=108)	1 h IGT (N=66)	p-value
Gestational age at birth, week, mean (SD)	38.7 (1.9)	38.5 (2.1)	38.6 (1.6)	NR
Pre-term birth	NR	NR	NR	NR
Pregnancy complications, %				
Pre-eclampsia	1.8	0.9	0	NR
Gestational hypertension	1.9	2.8	4.6	NR

Neonatal outcomes according to maternal glycaemic status

Adverse neonatal outcomes

Outcome, n (%)	Normal glucose tolerance (N=4,512)	h IGT (N=108)	h IGT (N=66)	p-value
Perinatal mortality	NR	NR	NR	NR
Mode of birth	NR	NR	NR	NR
Macrosomia	0.7	0	0	NR
LGA	6.4	5.6	13.6	p<0.05 1h IGT vs normal glucose
Birth injury	NR	NR	NR	NR
Neonatal hypoglycaemia	NR	NR	NR	NR
Admission to NICU	NR	NR	NR	NR
Long-term outcomes	NR	NR	NR	NR

RR or OR of outcomes in each glucose category relative to baseline category

LGA was more frequent in the 1 h IGT than normal glucose when adjusted for age, overweight, previous GDM, and a family history of diabetes (aOR 2.22; 95%CI 1.04 to 4.35, p=0.039). This was not significant in the 2 h IGT vs normal glucose group: aOR 0.75; 95% CI 0.33 to 1.72, p=0.508

Authors' Conclusions	Gestational IGT, defined as isolated hyperglycaemia, shows a metabolically heterogeneous phenotype in relation to the timing of isolated hyperglycaemia on the diagnostic OGTT

Abbreviations: aOR, adjusted odds ratio; BMI, body mass index; GDM, gestational diabetes; IGT, impaired glucose total; LGA, large-for-gestational age; NICU, neonatal intensive care unit; NR, not reported; OGTT, oral glucose tolerance test; SD, standard deviation; SGA, small-for-gestational age.

Study Design	Design Prospective cohort study Objective To evaluate the association between urinary phthalate metabolite concentrations on higher infant birth weight, stratifying by gradations of maternal glucose levels. Dates to 2008 Country US Setting Brigham and Women's Hospital, Boston

Table 64: Noor 2019

Study Reference **Noor 2019 (LIFECODES)**

Patient recruitment and eligibility

Recruitment: As part of the LIFECODES pregnancy cohort, women were recruited during their first prenatal visit (median 9.9 weeks gestation). **Inclusion criteria:** For this analysis, a subset of the population, who had available data on urinary phthalate metabolite concentrations, as a part of a nested case-control study among women who delivered between 2006 and 2008, with infants born ≥37 weeks of gestation, were included. **Exclusion criteria:** Women with a clinical diagnosis of GDM. **Other:** NR

Sample size

N screened/invited = NR
N eligible = 350 term births
N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR

ics N completed = NR

N excluded from analysis = 24 due to clinical diagnosis of GDM; 49 women with missing information on glucose levels N included in analysis = 277

Maternal demographics

Characteristic	Glucose <120 mg/dL (n=198)	Glucose 120 to <140 mg/dL (n=47)	Glucose ≥140 mg/dL without GDM (n=32)
Mean maternal age, years (SD)	31.3 (5.6)	31.6 (5.4)	33.7 (5.7)
Pre-pregnant BMI, kg/m^2	NR	NR	NR
BMI, kg/m^2	25.3 (5.3)	27.0 (6.1)	27.2 (5.2)
Weight, kg	NR	NR	NR
Non-Hispanic white	115 (58)	28 (60)	16 (50)

rence

Noor 2019 (LIFECODES)

Non-Hispanic black	38 (19)	7 (15)	3 (9)
Non-Hispanic Asian	4 (2)	4 (8)	4 (13)
Hispanic	29 (15)	6 (13)	7 (22)
Other	12 (6)	2 (4)	2 (6)
Medical history, n (%)	NR	NR	NR
Obstetric history, n (%)	NR	NR	NR
Education level			
<College	62 (32)	18 (39)	8 (27)
≥College	133 (68)	29 (61)	22 (73)

Maternal glycaemic characteristics
NR

Duration of follow-up
NR

Method of blood glucose measurement
Information on maternal glucose was collected from a non-fasting 50-g glucose load test in the second trimester as part of standard clinical screening for GDM undertaken by all study participants.

Diagnostic criteria and test(s)

Two step screening approach was taken for GDM diagnosis, with all women sitting for the 50-g, non-fasting GLT as the first step. Those women with glucose levels from the GLT ≥140 mg/dL were referred for further testing with a fasting 100-g 3 h OGTT. GDM was clinically diagnosed when a woman had two abnormal values from the 3 h OGTT following the elevated glucose value from the GLT. The Carpenter-Coustan criteria were utilised at the study institution.

Threshold cut-offs
>95 mg/dL (fasting); > 180 mg /dL (1h); >155 mg/dL (2h); > 140 mg/dL (3h)

In the analysis, maternal glucose was assessed as a categorial variable. Based on the GLT, glucose levels were classified as <120 mg/dL, 120– <140 mg/dL, and ≥140mg/dL to account for graduations of maternal glucose intolerance

Outcomes
LGA

Study Reference				
Adverse neonatal outcomes	Pregnancy outcomes NR Neonatal outcomes according to maternal glycaemic status			
	Outcome, n (%)	**Glucose <120 mg/dL (n=198)**	**Glucose 120 to <140 mg/dL (n=47)**	**Glucose ≥140 mg/dL without GDM (n=32)**
	Perinatal mortality	NR	NR	NR
	Mode of birth	NR	NR	NR
	Macrosomia	NR	NR	NR
	LGA			
	Noor 2019 (LIFECODES)			
	No	185 (93)	42 (89)	25 (78)
	Yes	13 (7)	5 (11)	7 (22)
	Birth injury	NR	NR	NR
	Neonatal hypoglycaemia	NR	NR	NR
	Admission to NICU	NR	NR	NR
	Long-term outcomes	NR	NR	NR
Authors' Conclusions	No conclusions related to the association between glucose levels and LGA are presented			

Abbreviations: BMI, body mass index; GLT, glucose level test; GDM, gestational diabetes; LGA, large-for-gestational age; NICU, neonatal intensive care unit; NR, not reported; OGTT, oral glucose tolerance test; US, United States.

Table 65. HAPO – Belfast (Thaware 2015) **Study Reference** **HAPO – Belfast (Thaware 2015)**

Population	Patient recruitment and eligibility
Characteristics	**Recruitment:** The HAPO study was a multicentre observational study that was designed to examine the associations between hyperglycaemia during pregnancy (short of diabetes) and adverse pregnancy outcomes. The Belfast HAPO Follow-Up Study is an ancillary study that is following up the offspring at the Belfast centre and represents a relatively unique cohort of carefully characterised subjects drawn from a homogenous population. Women participating in the Belfast centre of the HAPO study along with their offspring were invited for further follow-up examinations. **Inclusion criteria:** Offspring at age 5–7 years of women who had remained blinded to their HAPO study pregnancy OGTT results were included.

Study Reference

Study Design

Design
Cohort observational study

Objective
To examine the association of hyperglycaemia during pregnancy and anthropometry in 5 to 7 year old offspring whose mothers participated in the Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study at the Belfast Centre

Dates
NR

Country
Northern Ireland, UK

Setting

Exclusion criteria: NR **Other:** NR

Sample size
N screened/invited = 1,677 N eligible = NR

HAPO – Belfast (Thaware 2015) N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = NR
N excluded from analysis = NR N included in analysis = 1,320

Characteristic	Follow-up Status		p-value
	Yes (N=1,320)	No (N=284)	
Mean maternal age, years (SD)	30.04 (5.4)	28.3 (5.7)	<0.001

Maternal demographics

Study Reference

Cardiometabolic health			
Mean BMI at OGTT, kg/m^2 (SD)	28.1 (4.4)	29.0 (5.5)	0.02
BMI category at OGTT, kg/m^2, n (%)			
<28.5	809 (61.3)	154 (54.4)	0.007
28.5 to 32.9	345 (26.2)	74 (26.2)	
≥33	165 (12.5)	55 (19.4)	
Weight, kg			
Ethnicity, n (%)	NR	NR	NR
Medical history/risk factors, n (%)			
Hypertension	NR	NR	NR
Diabetes	NR	NR	NR
Smoking during pregnancy	280 (21.2)	105 (37.1)	<0.001
Alcohol use during pregnancy	359 (27.2)	71 (25.1)	0.47
Obstetric history, n (%)			
Primiparous	658 (49.8)	132 (46.6)	0.33
Parous without GDM	NR	NR	NR
Parous with GDM	NR	NR	NR
Mean education, years (SD)	15.0 (2.8)	14.1 (2.6)	<0.001

Maternal glycaemic characteristics

Pregnancy glucose tolerance	Follow-up Status	p-value

rence

	Yes (N=1,320)	No (N=284)	
FPG, mmol/L	4.6 (0.3)	4.6 (0.3)	0.39
75 g OGCT, mmol/L			
1 hour	7.4 (1.6)	7.4 (1.6)	0.42
2 hours	6.0 (1.1)	6.1 (1.2)	0.58
AUC PG	12.8 (2.0)	12.7 (2.1)	0.68

Study Reference HAPO – Belfast (Thaware 2015)

Methods

Duration of follow-up
NR

Method of blood glucose measurement
Participating women underwent a 75 g oral glucose tolerance

Diagnostic criteria and test(s)
- FPG
- 75 g OGTT

Outcomes
Primary endpoint
The relation between offspring adiposity at 5–7 years of age and maternal glycaemia was examined using continuous variables via regression analysis

Adverse neonatal outcomes

Pregnancy outcomes **outcomes** NR

Neonatal outcomes according to maternal glycaemic status

Offspring outcomes	Follow-up (N=1,320)
Mean BMI at follow-up, kg/m^2 (SD)	16.4 (1.9)
Mean BMI *z* score at follow-up (SD)	0.43 (1.01)
BMI *z* score category, n (%)	
≥85th percentile	318 (24.1)
≥95th percentile	143 (10.9)
≥99th percentile	62 (4.7)
Sum of skinfolds at follow-up, mm, geometric mean (range)	23.5 (18.6 to 28.2)

Outcome, OR (95% CI)	BMI *z* score			Sum of skinfolds ≥90th percentile (n=1,310)
	≥85th percentile (n=1,316)	≥95th percentile (n=1,316)	≥99th percentile (n=1,316)	
FPG				
Unadjusted	2.01 (1.37 to 2.96)*	2.37 (1.41 to 3.98)**	4.32 (2.07 to 9.04)*	2.48 (1.44 to 4.26)**
Adjusted model[a]	1.16 (0.76 to 1.76)	1.34 (0.76 to 2.35)	2.32 (1.05 to 5.13)	1.61 (0.90 to 2.89)
1-h PG unadjusted	1.06 (0.98 to 1.15)	1.01 (0.91 to 1.13)	1.06 (0.90 to 1.24)	1.02 (0.91 to 1.14)
2-h PG unadjusted	1.10 (0.99 to 1.23)	0.99 (0.85 to 1.15)	0.94 (0.75 to 1.18)	0.99 (0.84 to 1.16)

	AUC PG unadjusted	1.06 (1.00 to 1.13)	1.02 (0.93 to 1.11)	1.04 (0.92 to 1.18)	1.02 (0.93 to 1.12)
	GDM				
	Unadjusted	1.62 (1.17 to 2.25)	1.56 (1.01 to 2.41)	1.37 (0.72 to 2.63)	1.30 (0.81 to 2.09)
	Adjusted model[a]	1.18 (0.84 to 1.67)			

[a] Adjusted for maternal OGTT BMI and offspring birth weight *z* score
* P<0.001; **P<0.01

Study Reference	**HAPO – Belfast (Thaware 2015)**

Authors' Conclusions

Abbreviations: BMI, body mass index; FPG, fasting plasma glucose; GDM, gestational diabetes; HAPO, Hyperglycaemia and Adverse Pregnancy Outcome; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; PG, plasma glucose; SD, standard deviation; UK, United Kingdom.

Table 66. Verd 2016

Study Reference	**Verd 2016**
Study Design	Design Prospective cohort Objective To evaluate the association between mild gestational glucose tolerance impairment and the early cessation of exclusive breastfeeding. Dates January 2007 to December 2012 Country Spain Setting Majorca
	Patient recruitment and eligibility **Recruitment**: A population-based sample of mother-infant dyads attending a general care paediatric clinic in a middle class neighbourhood were enrolled. All mothers who attempted breastfeeding were invited to participate in "a study on infant feeding" upon their first well-child visit **Inclusion criteria**: Prenatal inclusion criteria were: (1) the routinely administered 24- to 28-week gestation 1 h OGTT, and (2) mothers had to be free of GDM. Post-birth inclusion criteria were (3) delivery at term (37 weeks of gestation), and (4) the mother initiated breastfeeding as planned. **Exclusion criteria:** NR **Other:**

stics

Sample size
N screened/invited = NR
N eligible = NR
N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = NR
N excluded from analysis = NR
N included in analysis =768

Study Reference **Verd 2016**

Maternal demographics

Characteristic	NCT: 1 h plasma glucose <7.8 mmol/L (N=616)	MIGT: 10.6 mmol/L >1 h OGTT ≥ 7.8 mmol/L (N=152)	P-value
Age, years, mean (range)	33 (20 to 45)	33 (25 to 42)	0.064
Cardiometabolic health			
Pre-pregnant BMI, kg/m^2	NR	NR	NR
BMI, kg/m^2	NR	NR	NR
Gestational weight gain	12 (1 to 39)	12 (4 to 27)	0.84
Ethnicity, n (%)	NR	NR	NR
Medical history/risk factors, n (%)	NR	NR	NR
Obstetric history, %	NR	NR	NR
Parity			
1	64	65	1.0
>1	36	35	1.0
Education level	NR	NR	NR

	Maternal glycaemic characteristics
NR	
Methods	Duration of follow-up NR Method of blood glucose measurement According to the recommendations of the American College of Obstetrics and Gynecology (ACOG), standard practice in the study's setting involves universal screening for GDM in all pregnant women at 24–28 weeks' gestation by a non-fasting 1 h 50 g glucose challenge test. Patients testing positive for the 1 h OGTT (1 h plasma glucose 7.8 mmol/L) were asked to return for a 3 h 100 g oral glucose tolerance test (3 h OGTT) Diagnostic criteria and test(s) • 1 h OGTT and 3 h OGTT • GDM, requires at least two of the following on the 3 h OGTT: fasting glucose 5.8 mmol/L, 1 h glucose 10.6 mmol/L, 2 h glucose 9.2 mmol/L, or 3 h glucose 8.1 mmol/L Threshold cut-offs • Normal glucose tolerance (NGT), defined by normal 1 h OGTT results (1 h plasma glucose < 7.8 mmol/L) • Mild impairment of glucose tolerance (MIGT), defined by a single abnormal value greater than or equal to 7.8 mmol/L, but less than 10.6 mmol/L; • GDM, requires at least two of the following on the 3 h OGTT: fasting glucose 5.8 mmol/L, 1 h glucose 10.6 mmol/L, 2 h glucose 9.2 mmol/L, or 3 h glucose 8.1 mmol/L Outcomes Delivery type
Adverse neonatal outcomes	Pregnancy outcomes NR

Study Reference	
	Verd 2016

Neonatal outcomes according to maternal glycaemic status

Outcome, n (%)	NGT: 1-hr plasma glucose <7.8 mmol/L (N=616)	MIGT: 10.6 mmol/L >1hr OGTT ≥ 7.8 mmol/L (N=152)	p-value
Perinatal mortality	NR	NR	NR
Mode of birth			

Induction of labour	NR	NR	NR
Vaginal delivery (eutocic)	82%	18%	0.67
Instrumental delivery	80%	20%	
C-section	79%	21%	
Macrosomia	NR	NR	NR
LGA	NR	NR	NR
Birth injury	NR	NR	NR
Birth weight, median g (range)	3272 (1995 to 4800)	3395 (2050 to 4390)	0.018
Neonatal hypoglycaemia	NR	NR	NR
Admission to NICU	NR	NR	NR
Long-term outcomes	NR	NR	NR

Authors' Conclusions	NA (not relevant to review objectives)

Abbreviations: ACOG, American College of Obstetrics and Gynecology; BMI, body mass index; GDM, gestational diabetes; LGA, large-for-gestational age; MIGT, midimpaired glucose tolerance; NA, not available; NICU, neonatal intensive care unit; NGT, normal glucose tolerance; NR, not reported; OGTT, oral glucose tolerance test.

Question 2: What are the most effective screening tests or strategies to identify women at risk of hyperglycaemia in pregnancy or GDM?

Table 67: Farrar 2016 (1678) Chapter 5.1 and Farrar 2017 (1675)

Study Reference	
Study Reference	**Farrar 2016 (1678) Chapter 5.1 and Farrar 2017 (1675)**
Study Design	Design Pooled analysis of individual participant data cohorts Objective To investigate whether multiple risk factor screening strategies represent a useful approach to screening for GDM. Dates September 2013 Country UK (BiB Cohort) and Ireland (ATLANTIC-DIP Cohort) Setting Bradford Royal Infirmary (BiB Cohort) and participating hospitals in the south-west of Ireland (ATLANTIC-DIP Cohort)
	Patient recruitment and eligibility **Recruitment:** BiB Cohort: All women planning to give birth at the Bradford Royal Infirmary were offered a 75 g OGTT (irrespective of risk factors) ATLANTICDIP: All women at participating hospitals were offered a 75 g OGTT (irrespective of risk factors) **Inclusion criteria:** NR **Exclusion criteria:** NR **Population Characteristics** **Other:** Uptake of 75 g OGTT offer varied between the two cohorts between 63% (BiB Cohort) and 58% (ATLANTIC-DIP Cohort). Sample size N screened/invited = NR N eligible = NR N enrolled = NR N excluded (with reason) = NR N lost to follow-up = NR N completed = 16,537 (10,432 from BiB Cohort and 6105 from ATLANTIC-DIP Cohort) N excluded from analysis = NR N included in analysis = 14,103 women (9939 from BiB Cohort and 4164 from ATLANTIC-DIP Cohort) with complete data on all risk factors

Farrar 2016 (1678) Chapter 5.1 and Farrar 2017 (1675)

Methods

Maternal risk factors

The following characteristics were examined due to their association with a greater risk of GDM development:

- age, examined yearly from 20 to 40 years
- obesity, measured by BMI at every 1.0 kg/m^2 unit increase from 15.0 to 40.0 kg/m^2
- parity, coded as primiparous (first pregnancy) or multiparous (second or subsequent pregnancy)
- ethnicity, coded as white, SA or other
- family history of diabetes
- GDM in previous pregnancy
- macrosomic baby (≥ 4 kg) in previous pregnancy

For age and BMI combined, results were presented for age ≥25 years and ≥30 years and BMI of ≥25 kg/m^2 and ≥30 kg/m^2.

The majority of women were either of white European (BiB Cohort and ATLANTIC-DIP Cohort) or SA ethnicity (BiB Cohort).

Reference standard

GDM diagnosed according to modified WHO 1999 criteria (fasting glucose level of ≥6.1 mmol/l, 2-hour post-load glucose level of ≥7.8 mmol/l). Gestational period test administered not specified.

Measures of test accuracy

For each risk factor the following were calculated with their SEs and 95% CIs:

- Sensitivity: The proportion of women with GDM who had the risk factor (i.e. proportion of GDM cases correctly identified by the test)
- Specificity: The proportion of women without GDM who did not have the risk factor
- Positive rate: The proportion of women with the risk factor (i.e. proportion who would be offered an OGTT)

To investigate the screening potential of offering an OGTT to any woman who has at least one from a set of risk factors, risk factors listed above were considered, with age ≥25 years or ≥30 years, and BMI ≥25 kg/m^2 or ≥30 kg/m^2. For each of the 287 possible combinations of these risk factors it was calculated whether or not each woman had at least one of the risk factors, and the sensitivity, specificity and positive rate associated with having one or more risk factors were estimated.

Risk factors that were 'dominated' by others were removed from this set of 287 possible combinations. A screening test is dominated if there is at least one other 'test' with both higher sensitivity and specificity, which would be preferred to the dominated test. Sensitivity and positive rate for the remaining non-dominated tests were plotted in ROC space.

Screening based on a predicted risk of GDM was examined, similar to screening strategies used to identify those at risk of cardiovascular disease. A logistic regression model was fitted to the data from both cohorts, regressing GDM incidence against the risk factors. The resulting log ORs from this regression model were used to calculate a predicted risk of GDM for each woman in the data set. The sensitivity and positive rate for predicting GDM at each percentage point of risk from 1% to 80% was calculated and plotted in ROC space. The same analyses were conducted on the separate and pooled data sets for comparison.

Test Accuracy

GDM prevalence

ATLANTIC-DIP prevalence of GDM (%): 10.2

Outcomes

BiB prevalence of GDM (%): 8.1

Study Reference

Risk factors for GDM screening

Study Reference

Page

Farrar 2016 (1678) Chapter 5.1 and Farrar 2017 (1675)

Risk factor	Sensitivity (%)	Specificity (%)	Positive rate (%)
BiB Cohort			
Age ≥25 years, BMI ≥30 kg/m^2	90.4	28.7	72.7
Age ≥25 years, BMI ≥30 kg/m^2, prior GDM	90.4	28.6	72.8
Age ≥25 years, BMI ≥30 kg/m^2, diabetes	91.6	23.2	77.7
Age ≥25 years, BMI ≥30 kg/m^2, diabetes, prior GDM	91.6	23.1	77.7
Age ≥30 years, BMI ≥30 kg/m^2, non-white	94.3	21.3	79.8
Age ≥30 years, BMI ≥30 kg/m^2, non-white, prior GDM	94.3	21.3	79.9
Age ≥25 years, BMI ≥25 kg/m^2, diabetes	94.4	16.9	83.8
Age ≥25 years, BMI ≥25 kg/m^2, diabetes, prior GDM	90.4	28.7	72.7
ATLANTIC-DIP cohort			
BMI ≥25 kg/m^2, non-white	90.1	36.8	66
Age ≥30 years, BMI ≥30 kg/m^2	90.8	28.6	73.4
Age ≥30 years, BMI ≥30 kg/m^2, non-white	93.9	26	76
Cohorts combined			
Age ≥30 years, BMI ≥30 kg/m^2, diabetes	90.0	24.6	76.4
Age ≥30 years, BMI ≥25 kg/m^2, diabetes, prior GDM	90.3	24.6	76.5
BMI ≥25 kg/m^2, non-white	92.0	24.0	77.3
BMI ≥25 kg/m^2, non-white, prior GDM	92.1	24.0	77.3
Age ≥25 years, BMI ≥30 kg/m^2	93.2	23.3	78.0
Age ≥25 years, BMI ≥30 kg/m^2, prior GDM	93.2	23.3	78.1
Age ≥30 years, BMI ≥30 kg/m^2, non-white	94.1	22.7	78.7
Age ≥30 years, BMI ≥30 kg/m^2 Age ≥25 years, BMI ≥25 kg/m^2, non-white, prior GDM	94.1	22.7	78.7
	95.9	16.5	84.5
Age ≥25 years, BMI ≥25 kg/m^2, prior GDM	95.9	16.5	84.5
NICE guideline recommended risk factors	78.2	31.7	67.2

Screening performance of one or more risk factor for identifying GDM was also presented graphically in the publication.

Risk association models

Risk factor	BiB Cohort	ATLANTIC-DIP Cohort

Study Reference Study Reference					
		OR	**95% CI**	**OR**	**95% CI**
	Age (per year)	1.09	1.08–1.1	1.1	1.07–1.12
	BMI (per kg/m^2)	1.06	1.05–1.08	1.13	1.11–1.15
	Ethnicity (non-white)	2.32	1.90–2.83	5.16	3.85–6.91
	Multiparity	0.89	0.73–1.08	0.74	0.58–0.96
	Family history of diabetes	1.36	1.14–1.63	1.42	1.17–1.80
	Previous GDM	1.54	1.12–2.13	–	–
	Previous macrosomia	5.9	3.78–9.22	–	–
	Additional results presented graphically include sensitivity and positive rate when using a risk prediction model to predict GDM and screening performance using risk prediction compared with having one positive risk factor or using age alone.				
Authors' Conclusions	No single method of risk factor screening is better overall. Risk factor screening based on having one or more risk factors and methods based on risk prediction or scoring performed similarly, suggesting that if risk factor screening is to be used, the simpler approach of offering an OGTT if at least one risk factor is present may be preferable. The potential benefits of offering universal testing must be weighed against any adverse effects and costs. Taken in this context the most efficient method of identifying women with GDM is likely to differ between populations. For high-risk populations in which the majority of women have a risk factor, especially a BMI of >30 kg/m^2 or advanced maternal age ≥25 years, universal testing may be most beneficial. For a young population of women with few risk factors, selective testing may be best; the use of risk factors in this population could be used to identify those at low risk who do not need testing and those remaining would be therefore offered an OGTT.				

Abbreviations: BiB: Born in Bradford; BMI: body mass index; CI: confidence interval; GDM, gestational diabetes mellitus; NICE: National Institute for Health and Care Excellence; NR, not reported; OGTT, oral glucose tolerance test; SE: standard error; SLR: systematic literature review; WHO: World Health Organization

Table 68: Farrar 2016 (1678) Chapter 5.2 and Farrar 2017 (1675)

Farrar 2016 (1678) Chapter 5.1 and Farrar 2017 (1675)

Study Reference **Farrar 2016 (1678) Chapter 5.2 and Farrar 2017 (1675)**

Study Design

Design
Systematic literature review
Objective
To investigate whether multiple risk factor screening strategies represent a useful approach to screening for GDM.
Search dates
6th June 2014, updated in August 2016
Country
Various
Setting
NA

Population Characteristics

Study eligibility

Inclusion (PICOS)

Population	Pregnant women without pre-existing diabetes
Intervention	**Index test** Any screening test that measured the association or predictive value of the following risk factors: • age • obesity and/or BMI • ethnicity (where applicable to the UK) • parity

Characteristic	Details
Design	Observational studies (prospective and retrospective cohort studies)
Sample sizes	Not summarised
Setting and timing	Not summarised
Participants	Not summarised

Study Reference

Study Reference Farrar 2016 (1678) Chapter 5.2 and Farrar 2017 (1675)

	• previous GDM, macrosomia or other GDM-related morbidity • family history of diabetes Only risk factors that were likely to be recorded in medical records without the need for further measurement were considered. **Diagnostic test** A diagnostic test (usually 75 g or 100 g OGTT) to diagnose GDM by recognised diagnostic criteria, or with criteria reported in the paper.
Comparator	
Outcomes	Numbers of women with and without GDM, according to the results of the diagnostic test. Studies had to report one of the following: the number of women with each risk factor, the sensitivity and specificity (screening performance) of the risk factor to identify GDM, data from which those statistics could be calculated or the accuracy of combinations of risk factors such as, numbers of risk factors present, risk models or scores based or measuring multiple risk factors, or the use of guideline recommendations.
Study design	Published, unpublished and ongoing observational studies, cohort studies, case–control studies or cross-sectional studies in English.

Exclusion

- Studies only reporting on the following risk factors: OGCT, FPG, vitamin D and genetic factors or studies that focused solely on biochemical tests such as the 50 g OGCT
- Studies reporting only ethnicity outside the UK
- Studies not reporting on at least one of the risk factors listed above
- Studies reporting associations between risk factors and GDM incidence, but not considering multiple risk factor screening • Studies examining the effect of screening based on a single risk factor Flow of Studies (PRISMA)
- Database results: 5867 (Jun 2014), 7858 (Aug 2016)
- Records after duplicates removed: 3140 (Jun 2014), 3586 (Aug 2016)
- Hand-searches/other sources: 13 (Jun 2014), 13 (Aug 2016)
- Title/abstracts reviewed: 3153 (Jun 2014), 4285 (Aug 2016)
- Full-texts reviewed: 181 (Jun 2014), 225 (Aug 2016)
- Articles included in synthesis: 29* (Jun 2014), 29 (Aug 2016)

*Of these, only 24 studies had sufficient data to allow inclusion in the analyses.

Included study characteristics

Study Reference	Farrar 2016 (1678) Chapter 5.2 and Farrar 2017 (1675)
Diagnostic criteria for GDM	Variety of criteria used, including IADPSG, Carpenter and Coustan, NDDG, WHO, ADA or local guidelines
Interventions and comparisons	Studies analysed: 24 studies* (Jun 2014), 29 studies (Aug 2016) • the screening performance of existing guideline recommendations (NICE, ADA, ACOG, ADIPS, Irish, French): 6 studies (Jun 2014), 6 studies (Aug 2016) • the number of risk factors for each woman: 7 studies (Jun 2014), 8 studies (Aug 2016) • a risk prediction model or a risk score: 6 studies (Jun 2014), 6 studies (Aug 2016) • various risk factors: 5 studies (Jun 2014), 9 studies (Aug 2016) *5 studies identified in the original SLR (Jun 2014) considered multiple risk factor screening, but reported insufficient data to allow inclusion in the analyses
Outcomes	Screening performance: sensitivity and specificity; calculated
Funding	NR
Conflicts of interest	NR

Definition of GDM

As defined in the individual trial using OGTT

	BEDIP-N (Benhalima 2018b) **Linked studies: Benhalima 2014; Benhalima 2018a (ID 554); Benhalima 2019**
Methods	Searches Sources searched • MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid SP) • EMBASE (via Ovid SP) • Maternity and Infant Care database (via Ovid SP) • CENTRAL (via The Cochrane Library/Wiley Interscience) • Reference searches of included journal articles and related systematic reviews Screening and selection process Title and abstract screening and then full-text screening performed in duplicate by two reviewers, with disagreements resolved by consensus or by a third reviewer. Quality assessment No formal quality assessment process was planned or undertaken because of the lack of any validated quality assessment tool for screening studies, and the diversity of type of study included. Synthesis methods The following screening performance statistics were calculated from the data presented for each study: • sensitivity (proportion of GDM cases correctly identified as high risk by screening) • specificity (proportion of women without GDM correctly identified as low risk) • positive rate (proportion of women who would be offered an OGTT if the risk factor combinations were present) Statistics were plotted across studies in ROC space by plotting detection rate against positive rate. The general performance of risk factor screening was then summarised, and the conclusions of each study considered.
Study Reference	**Farrar 2016 (1678) Chapter 5.2 and Farrar 2017 (1675)**
Meta-analysis methods for pooling of screening studies (such as the hierarchical summary receiver operator curves model) were considered, but not performed because of the considerable diversity across studies in terms of screening strategies and included risk factors.	
Test Accuracy Outcomes	Estimates of sensitivity and positive rate are presented graphically for each included study and synthesis findings are summarised in narrative. Overall, identifying greater numbers of women with GDM required offering an OGTT to increasing numbers. Findings were consistent across both the original SLR and the update. However, there did not seem to be an obvious 'best' approach. Across all studies, 3 screening methods based on counting the number of risk factors demonstrated the highest sensitivity against the lowest percentage of women offered OGTT. To achieve a sensitivity of 90%, an estimated 45% of women would need to be offered OGTT. Across studies reporting on the performance of current screening guidelines, none demonstrated a clearly superior sensitivity and positive rate. However, data reported by Coustan demonstrated a substantially inferior sensitivity for the French guidelines. To achieve a sensitivity of 90%, an estimated 60 to 80% of women would need to be offered OGTT. The screening performance of guideline recommendations was moderate at best, with the exception of the ACOG guideline when applied to an Irish or Spanish population and the ADA guideline when applied to an Irish population. Results across studies evaluating risk prediction models or risk scores were reasonably consistent, with all points lying approximately on a common ROC curve, suggesting that no specific risk scoring method is superior to another. Increasing sensitivity reduced specificity, for example to identify 80% of

Study Reference	women with GDM (sensitivity of 80%) using a risk prediction model or risk score, between 30% and 58% of women would need to undergo an OGTT (depending which risk model is used); to achieve a sensitivity of over 90%, nearly all women would need to undergo an OGTT. The conclusions of the study authors were varied, with 11 favouring universal diagnostic testing and 10 supporting some form of maternal characteristic/risk factor screening (universal screening and selective testing). Eight of the study authors made no firm recommendations. Of those that investigated current screening guideline recommendations (eight studies), seven did not recommend risk factor screening, three favoured universal diagnostic testing and four were undecided.
Quality Assessment	All included studies were observational, consisting of a mix of prospective and retrospective cohort studies. All studies used an OGTT to diagnose GDM, and all specified the diagnostic criteria used. Criteria varied between studies, therefore, there are differences in the thresholds used to define GDM. All of the risk factors examined are simple observable maternal characteristics/risk factors; the assessment of whether or not a risk factor is present, therefore, is unlikely to be subject to substantial measurement or reporting error or bias. Studies were diverse in their included populations. This heterogeneity limits the ability to draw conclusions across studies and generalise findings.
Authors' Conclusions	Pre-diagnostic risk factor screening is a poor method for identifying women with GDM and no single method of risk factor screening is better overall. Risk factor screening based on having one or more risk factors and methods based on risk prediction or scoring performed similarly, suggesting that if risk factor screening is to be used, the simpler approach of offering an OGTT if at least one risk factor is present may be preferable.

Abbreviations: ACOG: American College of Obstetricians and Gynecologists; ADA: American Diabetic Association; ADIPS: Australasian Diabetes in Pregnancy Society; BMI: body mass index; FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; IADPSG: International Association of the Diabetes and Pregnancy Study Groups; LR+, positive likelihood ratio; LR–, negative likelihood ratio; NA: not applicable; NDDG: National Diabetes Data Group; NGT, normal glucose tolerance; NICE: National Institute for Health and Care Excellence; NR: not reported; OGTT: oral glucose tolerance test; ROC: receiver operator curve; SLR: systematic literature review; WHO, World Health Organization

BEDIP-N (Benhalima 2018b)
Linked studies: Benhalima 2014; Benhalima 2018a (ID 554); Benhalima 2019

Table 69: BEDIP-N, Benhalima 2018b

Study Design	Design Prospective cohort Objective To evaluate the use of a GCT as a universal screening tool in a 2-step approach with the use of the 75 g 2-hour OGTT with the IADPSG criteria only if the GCT is abnormal To evaluate the characteristics of women with GDM who would be missed using a GCT threshold of ≥7.2 mmol/L, and to determine whether a modified two-step screening strategy with a GCT ≥7.2 mmol/L and clinical risk factors could improve the diagnostic strategy while exposing as few women as possible to the burden of an OGTT Also aimed to evaluate the tolerance of the tests and which screening strategy women preferred Dates April 2014 to March 2017 Country Belgium Setting Multicentre; 2 university hospitals and 5 non-university centres
Population Characteristics	Patient recruitment and eligibility **Inclusion criteria:** Women aged between 18–45 years who presented for prenatal care between 6–14 weeks of pregnancy **Exclusion criteria:** Multiple pregnancy, diabetes, history of bariatric surgery, miscarriage, chronic medical condition, gastro-intestinal surgery changing the absorption of glucose, planned home delivery or in a centre not participating in study, participation in another study 90 days before start of this study, other reasons causing normal follow-up and treatment to not be possible. **Other:** NR Sample size N screened/invited = 6–14 weeks: 2006; 24–28 weeks: 1884 N who received both GCT and OGTT: 1811 N eligible = NA N enrolled = NA N excluded (with reason) = abnormal FPG at 6–14 weeks: 19 (impaired fasting glycaemia [n=17], diabetes [n=2]); stopped before screening ≥24 weeks: 106 (miscarriage [n=26], abortion due to congenital malformation [n=16], mandatory bed rest [n=4], preterm delivery [n=1], stopped at own request [n=37]
Study Reference	

N lost to follow-up = 22

BEDIP-N (Benhalima 2018b) **Study Reference** **Linked studies: Benhalima 2014; Benhalima 2018a (ID 554); Benhalima 2019**

N completed = 1884 screened at 24–28 weeks gestation
N excluded from analysis = NR N included
in analysis = NR

Maternal demographics

Characteristic	Overall	GDM (N=231)	NGT (N=1610)	p value
Mean age, years	30.8 ± 4.1	32.0 ± 4.7	30.6 ± 3.9	**<0.001**
Cardiometabolic health				
Pre-pregnant BMI, kg/m^2 ± SD	24.1 ± 4.7	25.8 ± 5.5	23.8 ± 4.4	**<0.001**
BMI at first visit, kg/m^2 ± SD	24.7 ± 4.7	26.6 ± 5.3	24.4 ± 4.5	**<0.001**
Waist circumference at first visit, cm ± SD	87.3 ± 11.7	91.2 ± 13.0	86.5 ± 10.9	**<0.001**
Ethnicity, n (%)				
Ethnic minorities	213 (10.7)	43 (18.9)	132 (8.2)	**<0.001**
Medical history/risk factors, n (%)				
Systolic hypertension at first visit	44 (2.2)	7 (3.0)	30 (1.9)	0.215
Systolic hypertension at time of the OGTT	23 (1.2)	7 (3.1)	15 (0.9)	**0.014**
Diastolic hypertension at first visit	39 (1.9)	8 (3.5)	26 (1.6)	0.063
Diastolic hypertension at time of the OGTT	13 (0.7)	4 (1.7)	9 (0.6)	0.068
First degree family history of diabetes	255 (13.1)	42 (18.7)	185 (11.8)	**0.005**
Pre-pregnant smoking	587 (29.5)	80 (35.1)	456 (28.4)	**0.043**
Smoking during pregnancy	75 (3.8)	13 (5.7)	52 (3.2)	0.082
Pre-pregnant alcohol use	NR	NR	NR	NR
Obstetric history, n (%)				
Nulliparous	NR	NR	NR	NR
Parous without GDM	NR	NR	NR	NR
*Parous with GDM**	90 (9.3)	36 (30.2)	40 (5.3)	**<0.001**
Education level (highest education), n (%)				

Primary school	24 (1.2)	6 (2.6)	15 (0.9)	0.387
Until 15 years	92 (4.6)	11 (4.8)	69 (4.3)	
High school	278 (13.9)	36 (17.0)	189 (12.2)	

Study Reference

Bachelors	806 (41.8)	81 (37.5)	675 (43.1)
Masters	684 (35.5)	77 (35.6)	568 (36.2)

* Calculated in the number of women with a previous pregnancy Maternal glycaemic characteristics

Glucose tolerance	Overall	GDM (N=231)	NGT (N=1610)	p value
FPG, mmol/L	4.3 (4.1–4.6)	4.7 (4.4–5.1)	4.3 (4.1–4.5)	**<0.001**
75 g OGCT, mmol/L				
1 hour	7.1 (6.0–8.3)	9.5 (8.5–10.3)	6.8 (5.9–7.8)	**<0.001**
2 hours	6.2 (5.3–7.2)	8.6 (7.5–9.1)	6.0 (5.1–6.9)	**<0.001**

	BEDIP-N (Benhalima 2018b) **Linked studies: Benhalima 2014; Benhalima 2018a (ID 554); Benhalima 2019**
Methods	Index test and comparator **Benhalima 2018b** Screening with GCT (threshold of ⩾7.2 mmol/L) combined with clinical risk factors (ethnic minority background, BMI ⩾30 kg/m^2, history of GDM) compared to 1-step 75 g OGTT **Benhalima 2018a** Screening with different GCT thresholds (seeing which thresholds would be needed to achieve different sensitivity/specificity rates) compared to 1-step 75 g OGTT **Benhalima 2019** Risk factors only: maternal age and BMI, with and without other clinical risk factors – not extracted because these measurements were collected at the first visit (6–14 weeks) i.e. predictive instead of diagnostic Reference standard **24–28 weeks gestation, universal 2-step screening:** 50 g GCT and 75 g OGTT (universal screening of women without diabetes [FPG ≥7.0 mmol/L] or prediabetes [FPG ≥5.5 mmol/L and ≤6.9 mmol/L], as measured by FPG at 6–14 weeks gestation) • GCT thresholds were not prespecified because GCT was not yet validated with the 2013 WHO criteria and the results of the GCT were not used to inform treatment decisions for patients. The threshold of GCT used in the study for women to go on to receive OGTT was ≥7.2 mmol/L • The reference standard for the diagnosis of GDM was the 75 g OGTT with the 2013 WHO criteria (FPG ≥5.1 mmol/L, 1-h glycaemia ≥10.0 mmol/L, 2-h glycaemia ≥8.5 mmol/L, with diagnosis of GDM if ≥1 value was abnormal) Other tests **24–28 weeks gestation, modified 2-step screening:** 75 g OGTT only (no 50 g GCT) for women with a risk factor (e.g. ethnic minority, BMI ⩾30 kg/m^2, history of GDM), with universal 2-step screening for women without a clinical risk factor – but this was not used as a reference standard, because for the sensitivity and specificity analyses of the GCT combined with risk factors, only women who had received both the GCT and OGTT were included in the analyses The authors also calculated the number of women with GDM who would be missed when using a GCT threshold of 7.2 mmol/L compared to the universal one-step approach with the 75 g OGTT Measures of test accuracy Sensitivity, specificity, LR+, LR–, positive post-test probability, negative post-test probability
Test Accuracy	GDM prevalence
Outcomes	Range depending on test/threshold – see tables

Study Reference

GCT using a threshold of 7.2 mmol/L combined with clinical risk factors compared to 1-step test (75 g OGTT)

Risk factor	Total no. of OGTTs needed with GCT and risk factors combined, % (n)	No. of OGTTs needed based on GCT, % (n)	No. of OGTTs needed based on risk factors, % (n)	GDM, % (n)	Sensitivity, % (95% CI), n/N	Specificity, % (95% CI), n/N	LR+ (95% CI)	LR– (95% CI)	Positive post-test probability, % (95% CI)	Negative post-test probability, % (95% CI)
Ethnic minority background	41.3 (749)	31.2 (566)	10.0 (182)	9.8 (178)	78.1 (72.1–83.3), 178/228	64.0 (61.6–66.3), 1282/1583	2.2 (2.0–2.4)	0.34 (0.27–0.44)	24.3 (20.0–28.7)	4.8 (4.0–5.9)
BMI ≥30 kg/m^2	40.9 (741)	30.7 (557)	10.1 (184)	9.7 (176)	77.2 (71.2–82.5), 176/228	64.3 (61.9–66.7), 1020/1583	2.2 (2.0–2.4)	0.35 (0.28–0.45)	24.3 (20.0–28.7)	5.0 (4.1–6.0)
History of GDM	36.7 (665)	32.5 (589)	4.2 (76)	9.3 (169)	74.1 (67.9–79.7), 169/228	68.7 (66.4–71.0), 1089/1583	2.4 (2.1–2.6)	0.38 (0.30–0.47)	26.0 (21.4–30.7)	5.3 (4.4–6.4)
Any of the 3 risk factors	47.6 (868)	25.5 (462)	22.1 (400)	10.4 (189)	82.9 (77.4–87.5), 189/228	57.5 (55.0–59.9), 911/1583	1.9 (1.8–2.1)	0.30 (0.22–0.40)	22.4 (18.4–26.5)	4.2 (3.4–5.2)

- Using any of the 3 risk factors, the proportion of women that would be missed would be reduced to 17.1% (n=39) • 52.6% of all OGTTs could still be avoided
- Diagnostic accuracy of the GCT was not influenced by season (p=0.54), time after last meal before GCT (p=0.26), or random glucose value before the GCT (p=0.73)
- The global interaction with time of testing during the day of the GCT was not significant (p=0.06) but it was more often positive (≥7.2 mmol/L) in the afternoon (after 12 noon, 40.6%) than the morning (before 12 noon, 28.1%, p<0.001)
- The positive post-test probability of GCT was highest when GCT was performed before 12 noon (33.4%, AUC 0.82 [0.77–0.86]) compared to after 12 noon (22.9%, AUC 0.74 [0.69–0.79])

GCT thresholds at which GDM could be diagnosed without proceeding to OGTT (to achieve different specificity values)

- Specificity 99.9%: 11.1 mmol/L, sensitivity 3.1% (8 women [0.4%] were above this threshold)
- Specificity 99.1%: 10.2 mmol/L, sensitivity 7.9% (33 women [1.8%] were above this threshold)

Negative	GCTs, % (n) (n)	of GDM, %	(95 % CI), n/N	(95 % CI), n/N	CI)	CI)	post -test probability, % (95% CI)	post - test probability, % (95% CI)

BEDIP-N (Benhalima 2018b)
Linked studies: Benhalima 2014; Benhalima 2018a (ID 554); Benhalima 2019

Benhalima 2018a

Sensitivity and specificity of the GCT across different thresholds

BEDIP-N (Benhalima 2018b)
Linked studies: Benhalima 2014; Benhalima 2018a (ID 554); Benhalima 2019

Study Reference

	≥140 mg/dL (7.8 mmol/L)	23.9 (441)	7.5 (136)	59.6 (53.0–66.1), 136/228	81.0 (79.0–82.9), 1282/1583	3.1 (2.7–3.6)	0.50 (0.42–0.58)	31.7 (26.1–37.5)	6.9 (5.7–8.2)
	≥135 mg/dL (7.5 mmol/L)	29.1 (537)	8.3 (151)	66.2 (59.7–72.3), 151/228	76.1 (73.9–78.1), 1204/1583	2.8 (2.4–3.1)	0.44 (0.37–0.53)	29.1 (23.9–34.3)	6.2 (5.1–7.4)
	≥130 mg/dL (7.2 mmol/L)	34.9 (645)	9.1 (165)	72.4 (66.1–78.1), 165/228	70.2 (67.9–72.4), 1111/1583	2.4 (2.2–2.7)	0.39 (0.32–0.49)	26.4 (21.7–31.2)	5.5 (5.6–6.6)
	≥125 mg/dL (6.9 mmol/L)	40.8 (754)	9.8 (177)	77.6 (71.7–82.9), 177/228	64.2 (61.8–66.5), 1016/1583	2.2 (2.0–2.4)	0.35 (0.27–0.45)	24.3 (20.0–28.7)	4.9 (4.0–6.0)
	≥120 mg/dL (6.7 mmol/L)	48.4 (895)	10.3 (187)	82.0 (76.4–86.8), 187/228	56.0 (53.5–58.4), 886/1583	1.9 (1.7–2.0)	0.32 (0.24–0.43)	21.6 (17.8–25.6)	4.5 (3.7–5.5)

Authors' Conclusions

Benhalima 2018b

- A modified 2-step screening strategy with the GCT and clinical risk factors increased the sensitivity to 82.9%, and 52.6% of all OGTTs could be avoided, compared to the 1-step approach. A modified 2-step screening strategy might therefore be a practical alternative to the universal one-step approach with 75 g OGTT

Benhalima 2018a

- The GCT has a moderate diagnostic accuracy in a universal 2-step screening strategy for GDM using the 2013 WHO criteria. A GCT threshold of 140 mg/dL (7.8 mmol/L) had only a sensitivity of 59.6% and can therefore not be recommended in a 2-step approach for GDM using the 2013 WHO criteria. To achieve sensitivity rates ⩾70%, the threshold of the GCT would need to be reduced to at least 130 mg/dL (7.2 mmol/L)

Abbreviations: FPG, fasting plasma glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; LR+, positive likelihood ratio; LR–, negative likelihood ratio; OGTT, oral glucose tolerance test; NA, not applicable; NGT, normal glucose tolerance; NR, not reported; WHO, World Health Organization

Table 70: Project Viva, Gingras 2018

Study Reference	Project Viva, Gingras 2018
Study Design	Design Prospective pre-birth cohort Objective To evaluate the predictive value of fructosamine in pregnancy for abnormal gestational glucose tolerance and to assess the associations of fructosamine with maternal postpartum glycaemic indices Dates Blood samples collected between 1999–2002. Fructosamine measured in 2016–2017 Country USA Setting

Study Reference	Project Viva, Gingras 2018
Population Characteristics	Single centre (Atrius Harvard Vanguard Medical Associates) with 2 study hospitals Patient recruitment and eligibility Mothers were initially recruited during their first prenatal visit (median 9.9 weeks of gestation) at a multi-speciality group practice in eastern Massachusetts between 1999 and 2002 **Inclusion criteria:** • Less than 22 weeks pregnant at the time of enrolment • Receiving prenatal care at one of the selected practices • Plan on delivering at one of 2 study hospitals • Be able to answer questionnaires in English **Exclusion criteria:** • Multiple gestation • Planning to move aware before delivery • Planning to terminate the pregnancy Sample size N recruited to Project Viva: 2670 N still enrolled at time of delivery with live singleton birth: 2128 N excluded (with reason): 640 (missing fructosamine measurement: 540; measurement from samples with gross or moderate haemolysis or lipaemia: 49; pregestational type 1 or type 2 diabetes: 6; missing glucose tolerance status in pregnancy: 21; missing covariates: 24) N included in glucose tolerance assessment and fructosamine measurement: 1488 Maternal demographics Population characteristics were reported separately for different percentiles of fructosamine

Study Reference							
		<180 (5th percentile) N=69	≥180, <200 (5–25th percentile) N=286	≥200, <222 (25–50th percentile) N=366	≥222, <256 (50–75th percentile) N=393	≥256, <312 (75–95th percentile) N=299	≥312 (95th percentile N=75
	Mean age, years (±SD)	31.6 ± 4.9	32.0 ± 5.0	32.1 ± 4.7	32.3 ± 5.2	32.2 ± 5.0	32.5 ± 5.4
	Cardiometabolic health						
	Pre-pregnancy BMI, kg/m² (±SD)	56.1 ± 5.2	25.6 ± 5.9	24.6 ± 5.3	24.6 ± 5.3	24.0 ± 4.6	23.9 ± 4.6
	Pre-pregnancy overweight/obese, n (%)	34 (49)	120 (42)	126 (34)	130 (33)	97 (32)	24 (32)
	First trimester weight gain, kg (±SD)	2.7 ± 3.3	2.9 ± 2.9	2.8 ± 2.6	3.1 ± 3.0	2.6 ± 2.5	2.7 ± 3.1
	Ethnicity*						
	White, n (%)	51 (74)	193 (67)	267 (73)	289 (74)	221 (74)	55 (73)
	Black, n (%)	7 (10)	44 (15)	54 (15)	56 (14)	46 (15)	6 (8)
	Hispanic, n (%)	5 (7)	22 (8)	14 (4)	26 (7)	16 (5)	3 (4)
	Asian, n (%)	1 (1)	18 (6)	22 (6)	8 (2)	10 (3)	6 (8)
	Other, n (%)	5 (7)	9 (3)	9 (2)	14 (4)	6 (2)	5 (7)
	Medical history/risk factors						
	Smoking during pregnancy: yes, n (%)	7 (10)	42 (15)	50 (14)	42 (11)	31 (10)	5 (7)
Project Viva, Gingras 2018	**Obstetric history**						
	Nulliparous, n (%)	37 (54)	119 (42)	179 (49)	202 (51)	149 (50)	37 (49)
	Education level (highest education)						
	Education ≥ College degree, n (%)	45 (65)	182 (64)	273 (75)	257 (65)	204 (68)	52 (69)
	Household income >USD 70,000/year, n (%)	43 (70)	153 (58)	223 (66)	232 (65)	165 (60)	50 (72)
	Marital status: married/cohabiting, n (%)	65 (94)	263 (92)	344 (94)	359 (91)	282 (94)	70 (93)

*Due to rounding, not all column %s add up to 100%

Compared to participants included in the analysis, those excluded (N=640) were:

- Younger: 31.1 vs 32.1 years
- Less likely to be white: 52 vs 72%
- Less likely to have a college degree: 56 vs 68%
- Less likely to have an annual household income >USD 70,000: 55 vs 63%
- Less likely to be married or cohabiting: 88 vs 93%
- More likely to be overweight or obese prior to pregnancy: 41 vs 36%
- More likely to have developed GDM: 7.6 vs 4.9%
- Had greater pre-pregnancy BMI: 25.3 vs 24.7 kg/m^2

No difference was found between included and excluded participants for nulliparity, smoking in pregnancy and 50 g GCT results

	Maternal glycaemic characteristics NR
Methods	Index test and comparator **Fructosamine in blood sample collected at 24–28 weeks gestation** • Blood samples were collected between 1999 and 2002 and stored in liquid nitrogen. Fructosamine was measured in 2016–2017 (previous studies have suggested stability of the assay in frozen samples for short- or long-term) • A colorimetric assay was performed using a Roche P Modular system with reagents and calibrators from Roche Reference standard **2-step approach: 50 g GCT and 3h 100 g OGTT** • Women underwent routine clinical screening for GDM at 26–28 weeks gestation using a random 50 g GCT with venous blood sampling 1 hour post-load. Women with a blood glucose value <140 mg/dL were considered as having NGT, whereas women with a blood glucose value **≥140 mg/dL** were referred for a fasting 3h 100 g OGTT • Abnormal OGTT values were defined as: 1. >95 mg/dL at baseline 2. >180 mg/dL at 1h 3. >155 mg/dL at 2h 4. >140 mg/dL at 3h • Women were categorised based on the number and type of normal/abnormal values as follows: • Isolated hyperglycaemia (IH): normal OGTT values but abnormal GCT • IGT: 1 abnormal OGTT value • GDM: ≥2 abnormal OGTT values

Study Reference

Project Viva, Gingras 2018

Measures of test accuracy

Sensitivity and specificity of fructosamine levels ≥50th percentile (222 µmol/L), ≥75th percentile (256 µmol/L) and ≥95th percentile (312 µmol/L) to detect GDM and abnormal OGTT results (IGT and GDM)

ROC curves were generated to examine the predictive value of fructosamine for GDM, macrosomia, caesarean section, LGA, SGA an d preterm birth (<37 weeks gestation) *[N.B. results have only been extracted for GDM]*

Test Accuracy Outcomes

GDM prevalence

73 of 1488 participants, equating to 4.9%

Sensitivity and specificity of second trimester fructosamine to detect GDM during pregnancy in N=1488 women from Project Viva

Test	Sensitivity, %	Specificity, %	PPV, %	NPV, %
Fructosamine ≥222 µmol/L (≥50th percentile)	54.8	48.6	5.2	95.4
Fructosamine ≥256 µmol/L (≥75th percentile)	26.0	74.9	5.1	95.2
Fructosamine ≥312 µmol/L (≥95th percentile)	6.9	95.1	6.7	95.2

The sensitivity to detect GDM was 54.8% with a fructosamine level above the 50th percentile (222 µmol/L) with a specificity of 48.6%. Sensitivity was lower with a fructosamine level above the 75th percentile (256 µmol/L; 26.0%) and above the 95th percentile (312 µmol/L; 6.9%), while specificity was higher (74.9% and 95.1%, respectively). The authors note that this was as expected

ROC curves showed poor predictive value of fructosamine for GDM (AUC = 0.52 [this value did not change after implementing IPW in a sensitivity analysis with the purpose of addressing the issue of missing baseline fructosamine measurements in 589 women])

Authors' Conclusions

- The authors found weak associations between second trimester glucose values from the GCT and fructosamine, and fructosamine was a poor predictor of abnormal gestational glucose tolerance. Plasma fructosamine did not show adequate predictive characteristics for detecting impaired glucose tolerance in pregnancy or for predicting postpartum glycaemic status
- Fructosamine had a poor sensitivity to detect abnormal gestational glucose tolerance, with no acceptable cut-point identified

Abbreviations: AUC, area under the curve; BMI, body mass index; GCT, glucose challenge test; GDM, gestational diabetes mellitus; IGT, impaired glucose tolerance; IH, isolated hyperglycaemia; IPW, inverse probability weighting; LGA, large for gestational age; NGT, normal glucose tolerance; NPV, negative predictive value; NR, not reported; OGTT, oral glucose tolerance test; PPV, positive predictive value; ROC, receiver operating characteristic; SD, standard deviation; SGA, small for gestational age

Iimura 2015

Table 71: Iimura 2015

Study Reference	**Iimura 2015**
UK NSC external review	

Study Reference	
Study Design	Design Prospective cohort
Population Characteristics	Objective To determine the diagnostic potential of plasma lipids and apolipoproteins in GDM Dates December 2010 – July 2011 Country Japan Setting Japanese Red Cross Medical Center Patient recruitment and eligibility **Inclusion criteria:** Women with no previous diabetes **Exclusion criteria:** FBG >126 mg/dL, $HbA1_c$ >6.5%, blood glucose measurement of >200 mg/dL after 2h OGTT **Other:** NR Sample size N enrolled: 1183 N received GCT (20–28 weeks gestation): 1161; N did not receive GCT: 22 (referred to receive OGTT directly based on previous history of GDM or large size foetus) N GCT (+): 319; N GCT (–): 842* N received OGTT: 266; N OGTT not performed: 53** N OGTT (+): 45 (biochemical data n = 24, lipoprotein data n = 24); N OGTT (–): 221 (biochemical data n = 152, lipoprotein data n = 151) *Participants with a negative GCT were not allowed to undergo OGTT for ethic reasons to avoid unnecessary stress to the pregnant women **Participants for whom first check was carried out at the study centre but deliveries were in other facilities providing no follow-up data Maternal demographics

Study Reference

Ilmura 2015

Characteristic	GDM (–) (i.e. NGT)		GDM (+)		P value
	n	Mean (±SD)	n	Mean (±SD)	
Mean age, years	220	34.6 (0.35)	45	35.5 (0.78)	0.260
Cardiometabolic health					
BMI, kg/m²	218	20.3 (0.19)	45	22.1 (0.53)	0.002
Ethnicity	NR				
Medical history/risk factors	NR				
Obstetric history	NR				
Education level (highest education)	NR				

Maternal glycaemic characteristics

Glucose tolerance	GDM (–) (i.e. NGT)		GDM (+)		P value
	n	Mean (±SD)	n	Mean (±SD)	
1h GCT, mg/dL	221	147.8 (1.11)	45	161.5 (3.65)	<0.001

Maternal lipid apolipoprotein concentration

Lipid apolipoprotein concentration	GDM (–) (i.e. NGT)		GDM (+)		P value
	n	Mean (±SD)	n	Mean (±SD)	
TG, mg/dL	152	172.0 (6.53)	24	196.3 (15.2)	0.151
TC, mg/dL	151	256.0 (3.62)	24	252.6 (8.05)	0.711
LDL-C, mg/dL	152	130.6 (3.07)	24	131.8 (6.86)	0.883
HDL-C, mg/dL	152	82.6 (1.26)	24	81.3 (2.30)	0.621
ApoA-I, mg/dL	152	215.9 (2.50)	24	223.5 (6.14)	0.258

Ilmura 2015

ApoB, mg/dL	152	126.0 (2.56)	24	127.3 (5.75)	0.840
ApoB48, mg/dL	152	2.58 (0.14)	24	2.91 (0.37)	0.412
ApoC-III, mg/dL	152	14.3 (0.31)	24	15.6 (0.90)	0.197

Study Reference

Methods

Index test and comparator

Plasma lipids and apolipoproteins

- The timepoint at which measurements were taken is NR. Concentrations were measured in fasting samples
- TC, TG, HDL-C and LDL-C were quantified using enzymatic colorimetric methods according to manufacturer's protocols
- Insulin and ApoB-48 were measured by using a chemiluminescent enzyme-immunoassay system
- ApoA-I, ApoB and ApoC-III were determined by turbidometric immunoassay kits according to manufacture's recommendations
- CRP was measured using the Nanopia CRP kit (Sekisui Medical Co. Ltd.)

Reference standard

20–28 weeks gestation, 2-step screening: 1h 50 g GCT and 75 g OGTT

- GCT threshold (GCT [+]) for going on to receive OGTT was >130 mg/dL (note that a lower threshold than recommended by the Japan Assessment of GDM Screening Trial Group [>140 mg/dL] was used in this study because of the relatively older age of this cohort [average 34 years])
- Women who were GCT (+) underwent a 10–12h fasting 75 g OGTT with venous blood samples drawn at 0, 0.5, 1 and 2 hours. In accordance with the IADPSG and the ADA criteria, GDM was defined as plasma glucose concentration that exceeded one of the following measurements:
  1. Baseline: 92 mg/dL
  2. 1h-post OGTT: 180 mg/dL
  3. 2h-post OGTT: 153 mg/dL

Measures of test accuracy

ROC curves and AUC

Test Accuracy Outcomes

GDM prevalence

53 of 1183 participants, equating to 4.1% Predictive ability of lipid and apolipoprotein markers (AUC)

Marker	AUC	95% CI
TG	0.624	0.490–0.759
ApoC-III	0.583	0.451–0.715
ApoB48	0.568	0.439–0.697

Ilmura 2015

ApoA-I	0.560	0.438–0.684
HDL-C	0.531	0.420–0.641
ApoB	0.519	0.391–0.648
TC	0.518	0.388–0.647
LDL-C	0.515	0.393–0.636

The authors report that "only the AUC of ApoC-III was not good as a predictor of GDM" but it is unclear why this conclusion was reached, given the AUC of almost all other biomarkers is lower.

Study Reference

Authors' Conclusions	• It had been reported that ApoC-III could be a potential biomarker in women at 16–20 weeks gestation who subsequently develop GDM. However, the authors state that their data do not suggest that lipid or lipoprotein parameters have sufficient predictive power for GDM. It is known that during the mid-phase of pregnancy, maternal energy metabolism switches to enhanced lipolysis, a change that leads to increased levels of circulating fatty acids. This functional metabolic adjustment appears to be a general phenomenon during pregnancy and is unrelated to the mild glucose abnormality observed between GDM (+) and GDM (–) subjects. • During pregnancy, women with dysfunctional glucose metabolism have an associated abnormal lipid metabolism that results in a lipoprotein metabolism unlike that experienced when they are not pregnant. Prediction of GDM using only the ApoC-III value is not easy; however, pregnant women with higher concentrations of ApoC-III might require more medical supervision, and the same was true for ApoB48 with respect to diet (meal) absorptive ability.

Abbreviations: ADA, American Diabetes Association; Apo, apolipoprotein; AUC, area under the curve; BMI, body mass index; CRP, C-reactive protein; FBG, fasting blood glucose; GCT, glucose challenge test; GDM, gestational diabetes mellitus; $HbA1_c$, glycated haemoglobin; HDL-C, high density lipoprotein C; IADPSG, International Association of Diabetes and Pregnancy Study Groups; LDL-C, low-density lipoprotein C; OGTT, oral glucose tolerance test; NGT, normal glucose tolerance; NR, not reported; ROC, receiver operating characteristic; SD, standard deviation; TC, total cholesterol; TG, triglycerides

Table 72: Khalafallah 2016

Study Reference	**Khalafallah 2016**
Study Design	Design Prospective cohort Objective To provide an objective assessment of the utility of HbA1c when used as a screening tool in pregnancy. A direct comparison of HbA1c levels with results of the OGTT in women, tested concurrently at the 24–28 gestational week, was undertaken

Study Reference Khalafallah 2016

Dates

September 2012 to July 2014

Country

Australia

Setting

Tertiary referral teaching hospital (Launceston General Hospital)

Patient recruitment and eligibility **Inclusion criteria:**

Sequential women who were ≥18 years old and presented for OGTT test at 24–28 weeks gestation

Exclusion criteria:

Twin pregnancies

Women with an early diagnosis of GDM (prior to 24 weeks gestation) **Other:**

NR

Population Characteristics

Sample size

N enrolled: 480

Maternal demographics

There was limited reporting of maternal demographics, and reporting was for the whole population, rather than

GDM vs no GDM

- Median gestational age: 26 weeks; mean: 25.7 weeks (SD 3.3)
- Ethnicity: 93% Caucasian, 4% Asian, 3% Aboriginal Maternal glycaemic characteristics

Glucose tolerance	Mean	SD
Fasting glucose level, mmol/L	4.37	0.46
1h 75 g OGTT, mmol/L	6.85	1.7
2h 75 g OGTT, mmol/L	5.84	1.45

Study Reference

Methods	Index test and comparator **HbA1c at 24–28 weeks gestation** • HbA1c was measured by immunoassay using the DCA 2000, which measures HbA1c standardised to the National Glycohemoglobin Standardization Program (NGSP) Reference standard **75 g OGTT at 24–28 weeks gestation (1-step)** • The test was performed after an overnight fast of 10 hours. A sample was collected at baseline, then the patient consumed a 75 g glucose load (75 g dextrose in 300 mL carbonated liquid) within 5 minutes of starting the drink. Subsequent blood samples were collected at 1 and 2h post-start of the dextrose drink

Khalafallah 2016

	• Glucose concentration was measured within 3h of collecting the sample • In alignment with the 2013 ADIPS consensus guidelines for the testing and diagnosis of GDM in Australia, GDM was defined as present if: 1. FBG was ≥5.1 mmol/L **or** 2. 1h GTT was ≥10.0 mmol/L **or** 3. 2h GTT was ≥8.5 mmol/L Measures of test accuracy ROC curve; sensitivity; specificity; predictive values; false-positive and false-negative rates All analyses performed with SAS (V9.3)

GDM prevalence

57 of 480 participants, equating to 11.9%

Predictive values for arbitrary cut-off values of HbA1c

Test Accuracy

HbA1c arbitrary cut-off, %	Sensitivity, %	Specificity, %	PPV, %	NPV, %
10	0	99.7	0	88.8
6.1	2	99.7	50	89
6	4.1	99.7	66.7	89.2
5.9	6.1	99.7	75	89.4
5.8	8.2	99.7	80	89.6
5.7	10.2	99.5	71.4	89.8
5.6	12.2	99	60	90
5.5	22.4	98.2	61.1	91
5.4	26.5	95.4	41.9	91.2
5.3	34.7	88.4	27.4	91.5
5.2	55.1	79.7	25.5	93.4
5.1	61.2	67.6	19.2	93.3
5	69.4	51.9	15.4	93.1
4.9	73.5	31.4	11.9	90.4
4.8	81.6	18	11.1	88.6
4.7	95.9	10	11.8	95.1
4.6	95.9	4.6	11.2	90

Outcomes

Authors' Conclusions

- Employing a cut-off value for HbA1c at 5.4% (36 mmol/mol) for detecting GDM showed NPV of 91% and specificity of 95%. Similar results could be achieved with HbA1c level >5.1% as a screening tool for GDM
- The high specificity and NPV may be useful as an initial screening test for GDM. This may result in significant reduction in the burden of testing on both patients and testing facility, staff and resources
- Further investigations are required to integrate HbA1c as a single non-fasting screening tool for GDM with optimisation of the cut-off value

Abbreviations: ADIPS, Australian Diabetes in Pregnancy Society; FBG, fasting blood glucose; GDM, gestational diabetes mellitus; HbA1c, glycated haemoglobin; NGSP, National Glycohemoglobin Standardization Program; NPV, negative predictive value; NR, not reported; OGTT, oral glucose tolerance test; PPV, positive predictive value; SD, standard deviation

Table 73: Kosus 2012

Study Reference	Kosus 2012
Study Design	Design Retrospective cohort Objective To find optimal 100 g 3-hour OGTT threshold levels with high sensitivity and specificity for diagnosis of GDM in Turkish pregnant women Dates Between January 2008 and December 2009 Country Turkey Setting Fatih University, Faculty of Medicine
	Patient recruitment and eligibility **Recruitment:** NR **Inclusion criteria:** Healthy pregnant women screened for GDM and delivered at Fatih University were taken into the study

Study Reference

Population

Characteristics

Exclusion criteria: Patients with potential diabetic pregnancy (one or more of the following: previous history of GDM, including fan history, previous infants of congenital anomalies, previous unexplained foetal loss, hypertension, glucosuria by urine strip, previous complications, multi-foetal pregnancies, delivery prior to 24 completed weeks of gestation) were excluded from the study **Other:**

:tory of abetic

Sample size
N screened/invited = NR
N eligible = NR
N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = NR
N excluded from analysis = NR
N included in analysis = 808

Maternal demographics

	All patients (n=808)	Normal blood glucose (n=NR)	False positive GCT (n=NR)	Impaired glucose tolerance (n=NR)	GDM (n=NR)	p value

Study Reference **Kosus 2012**

Age, median years (IQR)	28 (7)	27 (8)	29 (8)	29 (7)	30 (9)	<0.001 normal vs FP-GCT and GDM
Cardiometabolic health, median (IQR)						
BMI, kg/m²	26 (5)	25.9 (4.4)	25.6 (4.1)	25.2 (3.9)	27.0 (4.8)	0.177
Ethnicity, n (%)	NR	NR	NR	NR	NR	NR
Medical history/risk factors, n (%)	NR	NR	NR	NR	NR	NR
Obstetric history, median (IQR)						
Parous without GDM	1 (2)	1 (1)	1 (2)	1 (2)	1 (2)	0.077
Gravidity	2 (2)	2 (2)	2 (2)	2 (3)	2 (2)	0.007 normal vs FP-GCT
Education level	*NR*	*NR*	*NR*	*NR*	*NR*	*NR*

Maternal glycaemic characteristics

Glucose tolerance, median (IQR)	Normal blood glucose (n=NR)	False positive GCT (n=NR)	Impaired glucose tolerance (n=NR)	GDM (n=NR)	p-value
GCT, mg/dL	109 (19)	144.5 (22)	159 (23)	165 (41)	<0.001 between all groups
100 g OGTT, mg/dL					
Fasting	-	81 (12)	87.5 (18)	95 (18)	<0.001 FP-GCT vs IGT and GDM
1-hour	-	147 (25)	170 (49)	194 (33)	<0.001 all groups except normal
2-hour	-	126 (25)	143.5 (34)	171 (18)	<0.001 all groups except normal
3-hour	-	98 (33)	112 (42)	140 (51)	<0.001 GDM versus normal, FP-GCT and IGT

Methods

Index test/Comparator

Screening between 24 and 28 weeks of gestation using the 3-hour 100 g OGTT. Optimal OGTT cut-off values for Turkish population were calculated by ROC curve analysis. Cut-off value for each hour was determined separately by using 0, 1, 2 and 3-hour glucose levels obtained by 100 g OGTT, with glucose levels showing highest sensitivity and specificity for prediction of GDM selected as cut-off points.

Reference standard

Screening between 24 and 28 weeks of gestation using the 1-hour 50 g GCT. If GCT ≥ 130mg/dL, diagnosis was confirmed by a 3-hour 100 g OGTT analysed by CC criteria, as the gold standard for the diagnosis of GDM. An abnormal 3-h 100 g OGTT was defined as two or more serum glucose values that meet or exceed the standards of CC criteria (fasting ≥95, 1-hour ≥180, 2-hour ≥155 and 3-hour ≥140 mg/dL). Serum glucose ≥200 mg/dL after GCT were also accepted as diabetes and a 3-hour OGTT was not performed.

Measures of test accuracy

Study Reference	Kosus 2012
	Measures of test accuracy (sensitivity, specificity and AUC) were reported. ROC curve analysis was performed to find the optimal cut-off point.

GDM prevalence

The prevalence of GDM was 8.1% (66 cases) by C&C criteria and 15.7% (127 cases) by index test criteria. Comparison of test accuracy between different screening methods

Method of screening	Cut-off, mg/dL	Sensitivity (%)	Specificity (%)	AUC	p-value	95% CI
Fasting 100 g OGTT	82.5	82.1	52.2	0.752	<0.001	0.678–0.825
1-hour 100 g OGTT	171.5	83.6	80.1	0.894	<0.001	0.854–0.934
2-hour 100 g OGTT	151.5	88.1	87.6	0.911	<0.001	0.868–0.954
3-hour 100 g OGTT	111.5	74.6	60.2	0.782	<0.001	0.708–0.857

Method of screening	Cut-off, mg/dL	Sensitivity (%)	Specificity (%)
– Screening for Gestatio Fasting 100 g OGTT	64.5	98.5	1.2
	71.5	97.0	7.5
– Screening for Gestational Diabetes	78.5	88.1	31.7
	82.5	82.1	52.2
	92.5	61.2	83.2
	95.5	49.3	87.0
1-hour 100 g OGTT	145.5	100.0	37.3
	158.5	100.0	62.7
	161.5	95.5	65.8
	165.5	91.0	69.6
	175	82.1	82.6
	185	67.2	89.4
2-hour 100 g OGTT	125.5	97.0	38.5
	130.5	95.5	50.9
	136.5	89.6	63.4
	145.5	88.1	80.1
	152.5	86.6	88.2
	161.5	79.1	91.9
3-hour 100 g OGTT	95.5	88.1	42.2
	101.5	80.6	49.7
	111.5	74.6	60.2
	125.5	64.2	84.5
	135.5	53.7	92.5
	145.5	40.3	98.1

UK NSC external review

- ROC curve showing the relationship between sensitivity and (1 – specificity) was presented graphically in the publication for the proposed new cut-offs for fasting, 1-hour, 2-hour and 3-hour 100 g OGTT results

Study Reference	**Kosus 2012**

- Additionally, the overlap between GDM diagnoses by CC criteria, NDDG criteria and the new proposed cut-off values was explored in the publication, as well as the correlation between OGTT results and demographic variables

Authors' Conclusions	Ethnic differences, environmental factors and nutritional habits may affect development of GDM. Application of some pre-determined nomograms to all races and ethnic groups can lead errors. In the present study, for better detection of GDM in Turkish society, different OGTT cut-off values from the CC and NDDG criteria were determined. Although the new criteria have some similarity with CC, they also have some important differences, especially in terms of the 3-hour cut-off. More extensive studies are needed for routine application of these new criteria during clinical practice.

Abbreviations: AUC: area under the curve; BMI: body mass index; CC: Carpenter and Coustan; CI: confidence interval; FP-GCT: false positive – glucose challenge test; GDM, gestational diabetes mellitus; IGT: impaired glucose tolerance; IQR: interquartile range; NDDG: National Diabetes Data Group; NR, not reported; OGTT, oral glucose tolerance test; SE: standard error; SLR: systematic literature review; WHO, World Health Organization

Table 74: Maesa 2018

Study Reference	
Study Design	Design Retrospective cohort study Objective To establish the usefulness of FG as screening of GDM in a population with low prevalence to avoid OGTT in low-risk pregnant women Dates Between September 2014 and February 2017 Country Spain Setting A tertiary hospital
Population Characteristics	Patient recruitment and eligibility **Recruitment:** NR **Inclusion criteria:** All pregnant women with data of FG corresponding to the visit made between the 24th and 28th gestational weeks, as well as the GCT and OGTT values in those with a GCT ≥140 mg/dL. **Exclusion criteria:** NR **Other:** NR Sample size N screened/invited = NR N eligible = NR N enrolled = NR N excluded (with reason) = NR **Maesa 2018** N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = 6573 Maternal demographics

Characteristic	Normal (n=6310)	Glucose intolerant (n=171)	GDM (n=92)
Age, years	31.84	33.06	33.90
Cardiometabolic health	*NR*	*NR*	*NR*

Ethnicity, n (%)	*NR*	*NR*	*NR*
Medical history/risk factors, n (%)	*NR*	*NR*	*NR*
Obstetric history, n (%)	*NR*	*NR*	*NR*

Maternal glycaemic characteristics

FG, mg/dL (mean, SD)	72.30 (11.29)	78.70 (14.89)	78.56 (19.16)
GCT, mg/dL (mean, SD)		154.57 (26.79)	
GCT <140 mg/dL, n (%)		5,305 (80.71)	
GCT ≥140 mg/dL, n (%)		1,268 (19.29)	

Index test/Comparator

FG assessed between the 24th and 28th gestational weeks

Reference standard

Screening performed using a 50 g GCT, followed by a 100 g OGTT when glycemia in GCT was ≥140 mg/dL assessed between the 24 and 28 gestational weeks. The 100 g OGTT follows recommendations and cut-off points of NDDG: fasting: 105 mg/dL, 1-hour: 190 mg/dL, 2-hour: 165 mg/dL, **Methods** and 3-hour: 145 mg/dL, diagnosing GDM when two or more measures are equal to or greater than these limits.

Measures of test accuracy

Sensitivity, specificity, positive predictive value, negative predictive value and likelihood ratios were reported for different cut-off points. The ROC curve was constructed using FG data and GDM diagnoses, the AUC was determined with the 95% confidence interval.

Assuming a diagnostic strategy in which FG was used as a screening test to prevent low-risk GDM pregnant women from being subjected to GCT, the percentage of pregnant women who would avoid the test was calculated for different thresholds.

114.37 (27.08) 182.84 (30.51)

Study Reference

Maesa 2018

GDM prevalence

Of 6,573 pregnant women included in the study, 92 (1.4%) had two or more altered points, so they were diagnosed with GDM.

Comparison of test accuracy between different screening methods

Study Reference

FG cut-off point for screening	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Positive likelihood ratio	Negative likelihood ratio	Ruling out GDM, n (%)
mg/dL	97.8	1.3	1.39	97.62	0.99	1.69	(1.28)
mg/dL	95.7	4.8	1.41	98.73	1.01	0.90	(4.79)
mg/dL	91.3	10	1.42	98.79	1.01	0.87	(10.03)
mg/dL	76.1	43.2	1.86	99.22	1.34	0.55	2,819 (42.89)
mg/dL	40.2	81.5	2.99	98.97	2.17	0.73	5,335 (81.17)

- The (AU)ROC curve, plotted for FG as screening test for GDM was 0.633 (95% CI 0.569 – 0.696).

Authors' Conclusions	FG could be considered a screening test to be performed before GDM diagnostic strategy, which would be an important benefit, both for the health system, in many cases overloaded, and for pregnant women who do not have to suffer the disadvantages of GCT and OGTT.

Abbreviations: AUC: area under the curve; BMI: body mass index; FG, fasting glycaemia; GCT: glucose challenge test; GDM, gestational diabetes mellitus; NPV: negative predictive value; NR, not reported; OGTT, oral glucose tolerance test; PPV: positive predictive value; ROC: receiver operator curve;

Table 75: Ohara 2016

Study Reference	Ohara 2016
Study Design	Design Retrospective study Objective To assess the impact of hyperemesis gravidarum on GDM screening given the similar effect that hyperemesis and GDM have on the starvation state of cells Dates Between 1st October 2010 and 30th September 2013 Country Japan Setting Tsukuba University Hospital

Ohara 2016

Patient recruitment and eligibility

Recruitment: NR

Inclusion criteria: The control group included all women without hyperemesis gravidarum who delivered within the specified period

Exclusion criteria: Women diagnosed with diabetes mellitus before pregnancy were excluded

Sample size

N screened/invited = NR
N eligible = 2112
N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = NR
N excluded from analysis = NR
N included in analysis = 33 were included in the hyperemesis group; the remaining 2079 women were included in the control group

Population

Maternal demographics

Ethnicity, n (%)	*NR*
Medical history/risk factors, n (%)	*NR*
Obstetric history, n (%)	
Nulliparous	(54.9)
Multipara (GDM status unspecified)	(45.1)
Education level	*NR*

Characteristic	Control group (n=2079)
Age, years (mean ± SD)	32.4 ± 5.4
Cardiometabolic health	
BMI at delivery, kg/m^2 (mean ± SD)	25.9 ± 4.2

Study Reference	

Maternal glycaemic characteristics

Glucose **tolerance**	**Control** group (n=2079)
Positive 50 g GCT (> 7.8 mmol/L)	432 (21.7)

Methods	Index test/Comparator Screening for GDM in the second trimester was carried out using a 50 g GCT with a cut-off value of 7.8 mmol/L performed from 24–27 weeks of gestation. Reference standard If screening test was positive, a definitive diagnosis was obtained using a 75 g OGTT. GDM was diagnosed on the basis of universal criteria established by the International Association of Diabetes and Pregnancy Study Group: a plasma glucose level that met or exceeded the fasting cut-off value (5.1 mmol/L), the 1-h cut-off value (10.0 mmol/L), or the 2-h cut-off value (8.5 mmol/L). Measures of test accuracy
Study Reference	**Ohara 2016**
	The positive predictive value was calculated by dividing the number of patients diagnosed with GDM by the number of women who were screened as positive for GDM.
Test Accuracy Outcomes	GDM prevalence Of women screened for GDM in the second trimester, 185/1994 (9.3%) were diagnosed with GDM. Analyses were performed after exclusion of GDM cases diagnosed in the first trimester. Comparison of test accuracy between different screening methods The positive predictive value of GDM screening in the second trimester using a 50 g CGT was 42.8%. Analyses were performed after exclusion of GDM cases diagnosed in the first trimester.
Authors' Conclusions	Hyperemesis gravidarum affects the positive GDM screening rate in the first trimester, but not in the second trimester, possibly due to related glucose metabolism abnormalities.

Abbreviations: BMI: body mass index; GCT: glucose challenge test; GDM, gestational diabetes mellitus; NR, not reported; OGTT, oral glucose tolerance test; SD: standard deviation

Table 76: Pawelec 2009

Study Reference	**Pawelec 2009**
Study Design	Design A prospective study Objective To calculate the real cost and clinical advantages and disadvantages of using a glucometer (the stick method) instead of the enzymatic method in screening for GDM Dates Between 2006 and 2008 Country Poland Setting First Clinic of Gynaecology and Obstetrics, Wroclaw Medical University
Population	Patient recruitment and eligibility **Recruitment:** Outpatients and hospitalised women at the gynaecology and obstetrics clinic **Inclusion criteria:** NR **Characteristics** **Exclusion criteria:** NR Sample size N screened/invited = NR **Pawelec 2009** N eligible = NR N enrolled = NR N excluded (with reason) = NR N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = 202 Maternal demographics

Characteristic	All patients (n=202)
Age pregnant women with GDM, years (mean)	36.2
Age pregnant women without GDM, years (mean)	27.8
Cardiometabolic health	NR
Ethnicity, n (%)	NR

Study Reference

Medical history/risk factors, n (%)	NR
Obstetric history, n (%)	NR
Education level	NR

Maternal glycaemic characteristics

Glucose tolerance	Venous blood screening (n=202)*	Finger capillary blood screening (n=202)*
Positive screening test (1-hour 50 g GCT >140 mg/dL), n (%)	7 (3.5)	9 (4.5)
Positive diagnostic test (2-hour 75 g OGTT >155 mg/dL), n (%)	6 (3.0) - the same patients were diagnosed as having GDM in both groups	

*Both venous and finger capillary blood screening were carried out in the same cohort of 202 patients.

Index test/Comparator

1-hour 50 g GCT carried out between 24–28 weeks of gestation using a finger capillary blood sample. If capillary blood glucose level of >140 mg/dL, GDM was investigated using a one-step 2-hour 75 g OGTT (cut-off 155 mg/dL [8.6 mmol/L]).

Reference standard

1-hour 50 g GCT carried out between 24–28 weeks of gestation using a venous blood. If venous blood glucose level of >140 mg/dL, GDM was investigated using the 2-hour 75 g OGTT (cut-off 155 mg/dL [8.6 mmol/L]).

Measures of test accuracy

Specificity and positive predictive value reported.

rence
09

GDM prevalence

Of the 202 patients in the study, 6 (3.0%) were diagnosed with GDM.

Method of screening	Specificity (%)	PPV (%)

Study Reference	
Accuracy Outcomes	Screening by finger capillary blood sample: 98.5*, 66.7* Comparison of test accuracy between different screening methods **Test**
	*Screening test accuracy, assuming the diagnostic test diagnosed 100% of GDM cases.
Authors' Conclusions	Screening using finger capillary glucose does not produce any harmful clinical effects because the diagnosis of GDM is done only after the diagnostic test, which is enzymatic. In short, both screening methods indicated the same patients as suspected GDM, and among these were the same patients later diagnosed as having GDM. The only disadvantage of the stick method is that because of the higher incidence of cases in which the result of the screening is not confirmed by the diagnostic test, there will be a few more women without GDM who will experience the stress connected with waiting for the results of the diagnostic test. However, the stick method also offers the advantage of a shorter screening time, which results in a shorter time to the diagnostic test if it proves necessary and, consequently earlier diagnosis and treatment. Another advantage of the stick method is the lower cost of nationwide screening. The money saved in this way can be directed to the treatment of other patients

Abbreviations: BMI: body mass index; GCT: glucose challenge test; GDM, gestational diabetes mellitus; NR, not reported; OGTT, oral glucose tolerance test; PPV: positive predictive value;

Table 77: Ryser Ruetschi 2016

Study Reference	Ryser Ruetschi 2016
Study Design	Design Cross-sectional study Objective To evaluate the sensitivity and specificity of a simplification for screening of gestational diabetes, where glucose loading would only be administered to women with fasting glycaemia between ≥4.4 and <5.1 mmol/L in the Swiss setting Dates Between October 2010 and April 2012 Country Switzerland Setting Various laboratories for glucose tolerance testing

Ryser Ruetschi 2016

Patient recruitment and eligibility

Recruitment: Collected anonymous oral glucose tolerance tests using from various laboratories
Inclusion criteria: To have a complete OGTT, with 3 values of glycaemia: fasting, after 1 and after 2 hours
Exclusion criteria: No exclusion criteria were applied

Sample size

N screened/invited = NR
N eligible = NR
N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = 2325
N excluded from analysis = 27 tests excluded as incomplete
N included in analysis = 2298

Maternal demographics

Population Characteristics

Characteristic	Patients with available demographics data (n=1932)
Age, years (mean, [range])	(15–50)
Cardiometabolic health	*NR*
Ethnicity, n (%)	*NR*
Medical history/risk factors, n (%)	*NR*
Obstetric history, n (%)	*NR*
Education level	*NR*

Maternal glycaemic characteristics

Glucose tolerance	Women with complete tests (n=2298)
Test results, n (% of all women) [95% CI]	
GDM	2047 (89.1) [9.7–12.3]
No GDM	251 (10.9) [NR]
Women with GDM, n (% of all women, % of women with GDM)	
Fasting glycaemia ≥5.1 mmol/L	119 (5.2, 47.4)
75g OGTT 2 hours ≥8.5 mmol/L	128 (5.6, 51.0)
75g OGTT 1 hour ≥10 mmol/L	103 (4.5, 41.0)
Women with or without GDM, (% of all women, % of women with GDM)	
Fasting glycaemia <4.4 mmol/L	1467[a] (63.8, NR)
Fasting glycaemia ≥4.4 – <5.1 mmol/L	712[b] (31.0, NR)

Study Reference

Total number of abnormal values across FPG and 75 g OGTT, n	
1 abnormal value	177
2 abnormal values	49

rence **Ryser Ruetschi 2016**

3 abnormal values	25
Fasting glycaemia as the only abnormal value, n (% of women with GDM)	73 women with GDM (29.1)

[a]Among these women, 54/1467 (3.7%) had GDM.
[b]Among these women, 78/712 (11%) had GDM.

Index test/Comparator

Fasting glycaemia separated in 3 categories: <4.4 mmol/L, ≥4.4 mmol/L but <5.1 mmol/L and ≥5.1 mmol/L, administered at 24 to 28 weeks of gestation.

Reference standard

1-hour and 2-hour 75 g OGTT using the IADPSG criteria (GDM diagnosed if fasting blood glucose ≥5.1 mmol/L, 1-hour blood glucose ≥10.0 mmol/L and 2-hour blood glucose ≥8.5 mmol/L) administered at 24–28 weeks of gestation.

Measures of test accuracy

Sensitivity, specificity and ROC curve reported.

Sensitivity and specificity calculated for two screening strategies: 1) stopping the test, avoiding the glucose loading and further glycaemia, if fasting glycaemia was <4.4 or ≥5.1 mmol/L and 2) excluding women with fasting glycaemia ≥5.1 mmol/L.

GDM prevalence

Among the 2298 complete tests, 251 women had GDM (10.9%; 95% CI 9.7–12.3%).

Comparison of test accuracy between different screening methods

UK NSC external revie

Test Accuracy Outcomes

Fasting glycaemia –cut Screening for Gestati-off (mmol/L)	Number of women at nal Diabet or above the uts -off, n (%)	Sensitivity, n women correctly diagnosed/all women with GDM (%)	Number of women avoiding the glucose overload, n (%)	Specificity, n women with GDM ruled out/all women without GDM (%)
4.0	(77.1%)	241/251 (96.0%)	(22.9%)	517/2047 (25.3%)
4.2	(56.1%)	223/251 (88.8%)	(43.9%)	980/2047 (47.9%)
4.4	(36.2%)	197/251 (78.5%)	(63.8%)	1413/2047 (69%)
4.6	(21.4%)	170/251 (67.7%)	(78.6%)	1725/2047 (84.3%)
5.1	(5.2%)	119/251 (47.4%)	(94.8%)	2179/2179 (100%)

- The sensitivity of avoiding the loading when fasting glycaemia was <4.4 or ≥5.1 mmol/l was 78.5% (95% CI 73.1–83.2). This strategy would avoid loading for 1586 (69.0%) women in the study population. Sensitivity was 76.1% in women <35 years old and 82.1% in women ≥35 years old.
- Excluding the 119 women with fasting glycaemia ≥5.1 mmol/l (i.e. meeting the criteria for GDM with fasting glycaemia only), the sensitivity of fasting glycaemia in the remaining 2179 women (132 with GDM) was 59.1% (95% CI 50.6–67.2).
- ROC curve showing the relationship between sensitivity and (1 – specificity) at various cut-offs of fasting glycaemia was presented graphically in the publication.

Study Reference	Ryser Ruetschi 2016
Authors' Conclusions	The results of the study demonstrate that, in a population with a low prevalence of GDM, using fasting glycaemia <4.4 mmol/l or ≥5.1 mmol/l to avoid glucose loading for GDM screening has a lower sensitivity than initially reported. Screening with fasting glycaemia is an attractive alternative to universal screening with the complete 75 g OGTT. In populations with a lower risk of GDM, however, screening using fasting glycaemia seems less sensitive, compared with settings with a higher prevalence of risk factors and/or of gestational diabetes. It appears that there is a correlation between the sensitivity of a strategy based on fasting glycaemia and the prevalence of gestational diabetes. Hence, it might be important to verify the sensitivity in a specific setting before implementing this simplified strategy.

Abbreviations: BMI: body mass index; GDM, gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; NR, not reported; OGTT, oral glucose tolerance test; ROC: receiver operator curve;

Table 78: Saeedi 2018

Study Reference	Saeedi 2018
Study Design	Design Cross-sectional Objective ***Primary aim:*** To evaluate the test characteristics of different levels of FBG values, traditional risk factors alone and in combination with RBG as indications to perform an OGTT for diagnosing GDM based on the modified IADPSG criteria, in a Swedish unselected population ***Secondary aim:*** To evaluate the test characteristics of the same factors in relation to the HAPO data OR 2.0 (model II) Dates 1 July 1994 – 30 June 1996 Country Sweden Setting Maternal health care in Örebro County, Sweden
	Patient recruitment and eligibility **Inclusion criteria:**

Study Reference **Saeedi 2018**

Population **Exclusion criteria: Characteristics** NR **Other:** NR

Sample size

N offered OGTT: 4918
N enrolled: 3616

Maternal demographics

Population characteristics were reported separately for women who underwent OGTT and no OGTT

Characteristic	OGTT (n=3616)	No OGTT (n=1302)	P value
Mean age, years ± SD	27.9 ± 4.8	28.5 ± 5.0	0.005
Cardiometabolic health			
BMI, kg/m^2 ± SD	23.8 ± 4.1	23.5 ± 3.8	0.18
Weight, kg ± SD	65.6 ± 12.1	64.9 ± 10.0	0.60
Height, cm ± SD	166 ± 6.0	166 ± 6.4	0.53
Ethnicity			
Non-Nordic origin, %	11.2	14.3	0.001
Medical history/risk factors			
Family history of diabetes (first degree relative), %	9.4	6.6	0.002
Prior GDM, %	1.3	0.5	0.020
Prior infant ≥4500 g, %	3.2	1.8	0.008
Obstetric history			
Primipara, %	46	30.6	<0.001
Education level (highest education)	NR		

Maternal glycaemic characteristics NR

Study Reference	Saeedi 2018
Methods	Index test and comparator **FBG, traditional risk factors and traditional risk factors in combination with RBG** • At the first maternal health visit, the traditional risk factors (first-degree relative, obesity [≥90 kg, pre-pregnancy weight], previous LGA infant [≥4500 g or ≥mean + 2SD] or GDM) and maternal characteristics (age, parity and ethnic origin) were recorded • RBG was measured 4 to 6 times during the pregnancy, starting at the end of the first trimester with ~6 week intervals. If any RBG measurements were ≥9.0 mmol/L, an OGTT was carried out immediately Reference standard Two models that were based on HAPO study outcomes were used (model I and model II) **Model I (modified IADPSG criteria) – 1-step** *(represents OR of 1.75 for adverse outcomes in the HAPO study)* • Fasting glucose ≥5.1 mmol/L and/or 2h 75 g OGTT ≥8.5 mmol/L **Model II – 1-step** *(represents OR of 2.0 for adverse outcomes in the HAPO study)* • Fasting glucose ≥5.3 mmol/L and/or 2h 75 g OGTT ≥9.0 mmol/L In accordance with WHO 1980 criteria, all women who attended maternal health care were offered a 75 g OGTT from gestational week 28–32. The women were instructed to intake carbohydrate rich food 2–3 days before the OGTT and fast after 10 pm the day before the test. Capillary glucose samples were taken at baseline (fasting) and 2h after the 75 g glucose load
	1h glucose test was not available and therefore was not included in the GDM diagnosis Whole blood capillary values were converted to plasma venous values by multiplying by a constant factor of 1.11 for fasting values and regarded as equivalent at 120 min. Random whole blood capillary value was not converted Measures of test accuracy Sensitivity, specificity, and predictive values were calculated using cross tabulations. ROCs of sensitivity plotted against 1-specificity were constructed for all possible diagnostic predicted venous fasting plasma glucose (pvFPG) cut-off values and the AUC was calculated. Comparisons were made using 95% CIs
	GDM prevalence • According to **model I** criteria (modified IADPSG criteria): 11.7% (10.3% on fasting alone, 2.7% on 2 h OGTT alone and 1.3% with both values elevated) • According to **model II** criteria: 7.2% (6.4% on fasting alone, 1.6% on 2 h OGTT and 0.8% with both values elevated) • 0.2% were diagnosed in early pregnancy using model I and II Characteristics of risk factors and tests for detecting GDM defined as model I (modified IADPSG criteria) or model II

rence
racy

Saeedi 2018	Test	Occurrence, % (n=3616)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	AUC, % (95% CI)
Model I[a]							

pvFPG threshold (equivalent FBG value)[b]						
≥4.4 (4.0) mmol/L	49	(94–98)	(55–59)	(21–25)	(98–99)	(75–78)
≥4.6 (4.1) mmol/L	41	(93–97)	(65–68)	(25–30)	(99–99)	(79–83)
≥4.8 (4.3) mmol/L	24	(88–94)	(83–86)	(41–48)	(98–99)	(86–90)
≥5.0 (4.5) mmol/L	14	(86–92)	(95–96)	(69–77)	(98–99)	(91–94)
Traditional risk factors[c]	16	(24–32)	(84–87)	(17–24)	(89–91)	(40–46)
Traditional risk factors[c] or RBG ≥8.0 mmol/L	19	(32–41)	(82–85)	(20–26)	(90–92)	(37–43)
Model II[d]						
pvFPG threshold (equivalent FBG value)[b]						
≥4.4 (4.0) mmol/L	49	(93–98)	(53–56)	(12–16)	(99–100)	(73–78)
≥4.6 (4.1) mmol/L	41	(93–98)	(62–65)	(15–19)	(99–100)	(78–82)
≥4.8 (4.3) mmol/L	24	(89–95)	(80–82)	(25–31)	(99–100)	(85–89)
≥5.0 (4.5) mmol/L	14	(87–94)	(91–93)	(42–50)	(99–100)	(89–94)
≥5.2 (4.7) mmol/L	8	(84–92)	(98–99)	(73–83)	(99–99)	(91–96)
Traditional risk factors[c]	16	(25–37)	(84–86)	(11–17)	(93–95)	(38–46)
Traditional risk factors[c] or RBG ≥8.0 mmol/L	19	(35–47)	(82–84)	(13–19)	(94–96)	(34–42)

[a]Model I (modified IADPSG criteria), 1.75 OR of adverse events in HAPO: equivalent cFBG ≥4.6 mmol/L or 2h OGTT ≥8.5 mmol/L [b]pvFPG and equivalent cFBG value using conversion factor of 1.11
[c]Model II, 2.0 OR of adverse events in HAPO: equivalent cFBG ≥4.8 mmol/L or 2h OGTT ≥ 9.0 mmol/L [d]Traditional risk factors = heredity (first-degree relative with diabetes), obesity (pre-pregnancy weight ≥90 kg), previous LGA infant (≥4500 g or ≥mean + 2SD), previous GDM

- The table shows that risk factor screening alone or in combination with random capillary glucose showed low sensitivity using both model I (28% and 36% respectively) and model II (31% and 41% respectively). Specificities for model I were 86% and 84% respectively and for model II were 85% and 83% respectively
- <u>For model I:</u>
- pvFPG cut-off values between 4.4 and 5.0 mmol/L had a sensitivity range between 89% and 96% and specificity between 57% and 96%

Study Reference	**Saeedi 2018**
	• The **optimal pvFPG cut-off value of 4.8 mmol/L** occurred in 24% of the patients with 91% sensitivity, 85% specificity and 88% AUC • For model II: • pvFPG cut-off values between 4.4 and 5.2 mmol/L had a sensitivity range between 89% and 96% and specificity between 54% and 98% • The **optimal pvFPG cut-off value of ≥5.0 mmol/L** occurred in 14% of the patients with 91% sensitivity, 92% specificity and 91% AUC
Authors' Conclusions	• In this cross-sectional, low-risk population-based study, current Swedish screening methods for GDM were found to be poorly predictive of GDM according to modified IADPSG criteria (model I, HAPO adverse event OR 1.75) and HAPO data (model II, OR 2.0) • • However, fasting glucose showed good test characteristics and could be an option for screening if resources with OGTT are limited pvFPG • cut-off values of 4.8 and 5.0 mmol/L, respectively, were the optimal criteria for referral for an OGTT A pvFPG of 4.8 and 5.0 mmol/L when using the model I and model II criteria would require 24% and 14% of women to progress to an OGTT, respectively. As the sensitivity increases for fasting glucose values, the specificity decreases. If the aim is to recognise disease, the sensitivity could be prioritised before specificity • Since the analysis was based on conversion of capillary blood glucose to venous plasma sample, there is a need for confirmation of the results

Abbreviations: AUC, area under the curve; BMI, body mass index; cFBG, capillary fasting blood glucose; CI, confidence interval; FBG, fasting blood glucose; GDM, gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; LGA, large for gestational age; NPV, negative predictive value; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; PPV, positive predictive value; pvFPG, predicted venous fasting plasma glucose; RBG, random blood glucose; ROC, receiver operating characteristic; SD, standard deviation; WHO, World Health Organization

Table 79: Temming 2016

Study Reference	**Temming 2016**
Study Design	Design Retrospective cohort study Objective To estimate if a threshold of a 1-hour GCT, alone or in combination with maternal risk factors, could achieve high enough specificity and positive predictive value to eliminate the need for a 3-hour GTT Dates Between 2004 and 2008

Study Reference Temming 2016

Country

US

Setting

University-based tertiary care centre (Barnes Jewish Hospital)

Patient recruitment and eligibility

Recruitment: All consecutive patients undergoing a 1-hour 50g GCT at Barnes Jewish Hospital, where a policy of universal GDM screening was applied. **Inclusion criteria:** Women were included in the study if they had a singleton gestation and completed 1-hour GCT testing followed by 3-hour GTT testing as appropriate after 20 weeks gestation.
Exclusion criteria: Women were excluded if there were no 3-hour GTT values available in the medical record or if they had Type I or Type II diabetes.

Other: NR

Sample size

N screened/invited = 6218
N eligible = 988
N enrolled = NR
N excluded (with reason) = 59 (treated as GDM) and 24 (result of 3-hour GTT recorded as “normal”, no numeric result) N lost to follow-up = 152
N completed = NR
N excluded from analysis = NR
N included in analysis = 753

Study Reference

Population Characteristics

Maternal demographics

Characteristic	1-hour GCT between 140mg/dL and 180mg/dL (n=657)	1-hour GCT ≥180 mg/dL (n=96)	Risk Ratio (95% CI)
Age, years (mean ± SD)	27.9 ± 6.3	28.7 ± 6.6	–
Age >30, n (%)	257 (39.1)	49 (51.0)	1.52 (1.05–2.21)
Cardiometabolic health			
Pre-pregnant BMI, kg/m²	*NR*	*NR*	*NR*
BMI, kg/m² (Mean ± SD)	34.10 ± 8.2	35.91 ± 8.6	–
Normal BMI, n (%)	213 (33.3)	27 (29.0)	0.84 (0.55–1.28)
Obese, n (%)	427 (66.7)	66 (71.0)	1.19 (0.78–1.81)
Morbidly obese, n (%)	129 (20.2)	23 (24.7)	1.25 (0.81–1.94)
Weight, kg	*NR*	*NR*	*NR*
Ethnicity, n (%)			
White	253 (33.6)		
African American	336 (51.1)	50 (52.1)	1.03 (0.71–1.50)
Asian (unspecified if South or East)	38 (5.0)		*NR*
Hispanic	60 (8.0)		*NR*
Temming 2016			
Other	7 (2.0)		*NR*
Mixed	*NR*	*NR*	*NR*
Medical history/risk factors, n (%)			
Chronic hypertension	37 (5.63)	8 (8.33)	1.43 (0.74–2.76)
Diabetes	*NR*	*NR*	*NR*
Tobacco use	105 (16.0)	18 (18.8)	1.18 (0.74–1.90)
Pre-pregnant smoking	*NR*	*NR*	*NR*
Pre-pregnant alcohol use	*NR*	*NR*	*NR*
Obstetric history, n (%)			
Nulliparous	*NR*	*NR*	*NR*
Parous without GDM	*NR*	*NR*	*NR*
Parous with GDM	144 (21.9)	53 (55.2)	3.48 (2.40–5.02)
Education level	*NR*	*NR*	*NR*

Study Reference

Methods

Index test/Comparator

Screening by 1-hour 50 g GCT between 24–28 weeks unless risk factors suggested need for earlier testing, although only those with testing performed after 20 weeks were included for this analysis. Risk factors leading to early testing included a history of previous GDM, obesity with BMI ≥30.0 kg/m^2, history of macrosomic infant in a prior pregnancy, first degree relative with diabetes mellitus, or glycosuria. For women with a normal early 1-hour GCT, screening was repeated between 24–28 weeks and only the second was included for analysis. An elevated 1-hout GCT was defined as ≥140 mg/dL and an extremely elevated 1-hour was defined as ≥180 mg/dL, as originally suggested by CC criteria.

Reference standard

Diagnostic testing with a 3-hour 100 g GTT completed generally within 1 week of an initial positive screening test (elevated 1-hour GCT ≥140 mg/dL).

GDM was diagnosed by having 2 or more abnormal values using NDDG criteria (fasting ≥105 mg/dL, 1-hour ≥190 mg/dL, 2-hour ≥165 mg/dL, 3-hour ≥145 mg/dL) or using more stringent CC criteria (fasting ≥95 mg/dL, 1-hour ≥180 mg/dL, 2-hour ≥155 mg/dL, 3-hour ≥140 mg/dL).

Measures of test accuracy

1-hour GCT results were categorized by 20 mg/dL increments between 160 mg/dL and 220 mg/dL. ROC curves, AUC, sensitivity, specificity, positive and negative predictive values were reported. The AUC was calculated for each of the thresholds between 160 mg/dL and 220 mg/dL for the 1-hour GCT to diagnose GDM using both CC and NDDG thresholds. The optimal cut-point was identified using the Youden index which maximizes the sum of sensitivity and specificity. Analysis was repeated for each of the thresholds amongst women with individual and combinations of specific risk factors, including maternal BMI ≥30 kg/m^2, history of GDM, and maternal age. Calculations were performed for both NDDG and CC criteria for diagnosis of GDM.

Test Accuracy Outcomes

GDM prevalence

Of 6218 women screened, 59 (0.95%) were treated as GDM without a 3-hour GTT based on provider preference.

Of the eligible women with an elevated 1-hour GCT, 165 women (2.7% of the total cohort n=6218) were diagnosed with GDM by NDDG criteria, and 250 (4.0% of the total cohort n=6218) were diagnosed with GDM by CC criteria. Comparison of test accuracy between different screening methods

Temming 2016

	NDDG Criteria for GDM				CC Criteria for GDM			
1-hour GCT value (mg/dL), n	**Sensitivity (%, 95% CI)**	**Specificity (%, 95% CI)**	**(AU)ROC curve (%, 95% CI)**	**PPV (%, 95% CI)**	**Sensitivity (%, 95% CI)**	**Specificity (%, 95% CI)**	**(AU)ROC curve (%, 95% CI)**	**PPV (%, 95% CI)**
≥160, n=283	65.5% (57.7–72.7)	70.2% (66.4–73.9)	0.678 (0.638–0.719)	38.2%	58.4% (52.0–64.6)	72.8% (68.6–76.6)	0.656 (0.620,0.692)	51.6% (45.6–57.5)
≥160 and history of GDM, n=115	66.7% (57.6–74.9)	55.4% (43.4–67.0)	0.610 (0.540–0.681)	(32.5–44.1) 71.3% (62.1–79.4)	59.2% (51.5–66.6)	47.8% (26.8–69.4)	0.535 (0.425–0.646)	89.6% (82.5–94.5)
≥160 and history of GDM, age ≥30, BMI ≥30 kg/m^2, n=40	65.3% (50.4–78.3)	65.2% (42.7–83.6)	0.653 (0.532–0.773)	80.0%	56.9% (44.0–69.2)	57.1% (18.4–90.1)	0.570 (0.363–0.777)	92.5% (79.6–98.4)

	≥180, n=96	30.3% (23.4–37.9)	92.2% (89.7–94.2)	0.612 (0.576–0.649)	(64.4–90.9) 52.1% (41.6–62.4)	24.8% (19.6–30.6)	93.2% (90.7–95.3)	0.590 (0.561–0.619	64.6% (54.2–74.1)
Study Reference	≥180, and history of GDM, n=53	(24.4–41.6) 38.8%	90.3) 82.6% (71.8–	(56.3– 0.635) 0.607 (0.502–	86.2) 82.6%	(21.1–34.9) 32.3%	92.5) 71.4% (29.0–	0.529 (0.437–0.622)	90.6% (79.3–96.9)
	≥180, and history of GDM, age ≥30, BMI ≥30 kg/m^2, n=23	(25.2–53.8) 14.5% (9.6–	90.3) 98.1% (96.7–	0.712) 0.563		(21.2–45.1) 11.2% (7.6–	96.3) 98.6% (97.2–	0.519 (0.329–0.708)	91.3% (72.0–98.9)
	≥200, n=35	20.9) 16.3%	99.1) 95.9% (88.6–	(0.536–0.591) 0.561	(61.2–95.0) 68.6% (50.7–83.1) 87.0% (66.4–97.2)	15.8) 12.6% (8.1–	99.4) 95.7% (78.1–	0.549 (0.529–0.569)	80.0% (63.1–91.6)
	≥200, and history of GDM, n=23	(10.2–24.0) 22.4%	99.2) 91.3% (72.0–	(0.521–0.601) 0.569	84.6% (54.6–	18.5) 18.5% (9.9–	99.9) 85.7% (42.1–	0.541 (0.492–0.591)	95.7% (78.1–99.9)
	≥200, and history of GDM, age ≥30, BMI ≥30 kg/m^2, n=13	(11.8–36.6) 4.9% (2.1–	98.9) 99.5% (98.5–	(0.485–0.652)	98.1)	30.0)	99.6) 99.6%	0.521 (0.373–0.669)	92.3% (64.0–99.8)
	≥220, n=11	9.3) 5.7% (2.3–	99.9) 100.0%	0.522 (0.538)	72.7% (39.0– 94.0)	3.6% (1.7– 6.7)	(98.6– 100.0)	(0.505– 0.516 (0.504–0.528)	81.8% (48.2–97.7)
	≥220, and history of GDM, n=7	11.4)	100.0)(95.1–	0.528 (0.508–0.549)	100.0% (59.0–100.0)	4.0% (1.6–8.1)	100.0% (85.2–100.0)	0.520 (0.505–0.535)	100.0% (59.0–100.0)
	≥220, and history of GDM, age ≥30, BMI ≥30 kg/m^2, n=4	8.2% (2.3–19.6)	100.0% 100.0)(85.2–	0.541	100.0%	6.2% (1.7–15.0)	100.0% (59.0–100.0)	0.531 (0.051–0.560)	100.0% (39.8–100.0)
		32.5%	82.4% (71.8–	0.575 (0.514– (61.7–	75.5% 27.6% 78.3%	(0.502–0.580)	(39.8–100.0)		

For 1-hour GCT using NDDG criteria, the (AU)ROC was 0.730. The Youden cut point was 157.5 mg/dL at (AU)ROC cut point 0.680. For 1-hour GCT using CC criteria, the (AU)ROC was 0.693. The Youden cut point was 158.5 mg/dL at (AU)ROC cut point 0.660.

Authors' Conclusions	Even with an extremely elevated 1-hour GCT result ≥ 200 mg/dL, 20%–33% of patients would be over diagnosed with GDM if the 3-hour GTT was omitted. Although the addition of maternal risk factors marginally improves the specificity and positive predictive value of an extremely elevated 1-hour, it would only eliminate the need for a 3-hour GTT in a few select patients, making this less practical. These findings support the need for a diagnostic 3hour GTT even in those patients with extremely elevated 1-hour results

Abbreviations: AU(ROC), area under the (receiver-operator curve); BMI, body mass index; CC, Carpenter and Coustan; CI, confidence interval; GCT; glucose challenge test; GDM, gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; PPV, positive predictive value; pvFPG, predicted venous fasting plasma glucose; RBG, random blood glucose; ROC, receiver operating characteristic; SD, standard deviation; WHO, World Health Organization

Table 80: Theriault 2014

Population Characteristics	Patient recruitment and eligibility **Recruitment:** Women recruited prospectively at their first prenatal visit **Inclusion criteria:** Women were eligible if they were at least 18 years old and without renal and hepatic disease **Exclusion criteria:** Pregestational diabetes (n=65), multiple pregnancy (n=107), uncertain diagnosis (absence of screening and/or diagnostic tests and gestational age at delivery unknown or before 32 WG, n=395) and delivery outside of study centres (n=91) **Other:** Data from 63 patients were removed from the databank at their request Sample size N screened/invited = 7929 N eligible = NR N enrolled = NR N excluded (with reason) = 721 (see reasons above) N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = 7208

Study Reference **Theriault 2014**

Study Design

Design
Cohort study
Objective
To validate the performance of proposed clinical risk-prediction models for identifying women who developed GDM and those who required insulin therapy in order to improve risk stratification and facilitate follow up and prevention Dates Between March 2005 and April 2010
Country
Canada
Setting
Databank

Study Reference **Theriault 2014**

Maternal demographics

	Women who developed GDM (n=381)	Women who did not develop GDM (n=6827)	p-value
Age, mean years (SD)	30.9 (4.6)	29.4 (4.3)	<0.001
Cardiometabolic health			
Pre-pregnant BMI, kg/m^2	28.3 (7.1)	24.1 (5.1)	<0.001
BMI at first prenatal visit, kg/m^2	29.3 (6.9)	25.1 (5.1)	<0.001
Weight, kg	*NR*	*NR*	
Ethnicity, %			
Caucasian	96.0	96.8	*NR*
Medical history/risk factors, %			
Family history of diabetes (1st degree)	35.2	16.5	<0.001
Family history of diabetes (1st or 2nd degree)	72.4	54.2	<0.001
Smoking (unspecified if pre-pregnancy)	14.7	10.0	<0.05

Obstetric history, %			
Nulliparous	42.4	47.3	*NR*
Parous with GDM	28.2	1.5	<0.001
Gestational hypertension or preeclampsia	8.3	3.9	<0.001
Macrosomic infant (≥4000 g)	14.2	6.0	<0.001
Recurrent spontaneous abortion (≥3)	1.8	1.9	*NR*
Foetal death ≥20 WG	1.2	0.6	*NR*
Education level	*NR*	*NR*	*NR*
Exercise before pregnancy, %			
0–3 times/month	32.4	24.7	<0.05
1–3 times/week	58.7	61.0	*NR*
≥4 times/week	9.0	14.3	<0.05
Family income ≥60,000 CDN$, %	47.5	55.2	<0.05
Need for insulin therapy (current pregnancy), %	66.4	–	*NR*

Study Reference **Theriault 2014**

Maternal glycaemic characteristics

- Among the 381 women with GDM who were included, 87 were diagnosed based on the result of the GCT (≥10.3 mmol/L) alone and 172 were diagnosed after the OGTT (≥2 values exceeding the thresholds of 5.3, 10.6 and 8.9 mmol/L at 0, 1 and 2 h, respectively)
- Information retrieved in the medical records was used to establish another GDM subgroup of 122 women who received insulin during pregnancy without undergoing an OGTT. This group consisted of patients for whom (1) the screening and diagnostic tests were either not performed, results were unavailable or borderline, and (2) frankly abnormal results on glucose monitoring led to the decision to start insulin therapy during pregnancy. This allowed to identify all women with severe GDM and mitigate the false negative rate of the GCT
- Finally, 151 women were diagnosed with impaired glucose tolerance (IGT) after the OGTT (1 value exceeding the thresholds)

Index test/Comparator
Four clinical risk-prediction models based on risk factors assessed at 24–28 WG using a self-administered questionnaire, anthropometric measurements and clinical

Methods

Model/study nar	Methodology	Clinical risk factors	Scoring system
	Cohort of 3131 women (113 GDM)	Maternal age	
	Naylor et al. (Canada)	Derivation: 1560 women (44 GDM) Validation: 1571 women (69 GDM) Diagnostic test: 100 g OGTT (all)	BMI before pregnancy Ethnicity	≤22.0 (0); 22.1–25.0 (2); ≥25.1 (3) White (0); Black (0); Asian (5); Other (2)
	Caliskan et al. (Turkey)	Cohort of 4612 women (143 GDM) Validation: 422 women (14 GDM) Screening with GCT Diagnostic test: 100 g OGTT (all)	Maternal age ≥25 years BMI before pregnancy ≥25 kg/m^2 Prior adverse obstetric outcome* Fam. history of diabetes (1st degree) Prior macrosomic foetus (>4000 g)	1 point for each risk factor
	van Leeuwen et al. (Netherlands)	Cohort of 995 women (24 GDM) Screening with GCT Diagnostic test: 75 g OGTT (not all) Prediction of missed GDM Probability of GDM calculated	BMI before pregnancy Ethnicity Fam. history of diabetes (1st or 2nd degree) Previous GDM	probability of GDM = 1/[1 + exp(-β)] β= [-6.1 + (0.83 x non-Caucasian ethnicity) + (0.57 x positive fam. history of diabetes) - (0.67 x multipara without history of GDM) + (0.5 x multipara with history of GDM) + (0.13 x BMI)]
	Teede et al. (Australia)	Cohort of 4276 women (356 GDM) Derivation: 2880 women (250 GDM)	Maternal age BMI at first visit	<25 (0); 25–34 (1); ≥35 (2) <20.0 (0); 20.0–34.9 (1); ≥35.0 (2)

		Validation: 1396 women (106 GDM) Screening with GCT Diagnostic test: 75 g OGTT (not all)	Ethnicity	Caucasian (0); African (1); Asian (0–1–2)**; Polynesian (1); Other (0)

Study Reference	**Theriault 2014**			
			Fam. history of diabetes (1st degree)	No (0); Yes (1)
			Past history of GDM	No (0); Yes (2)

*Prior obstetric outcome defined as recurrent spontaneous abortion (>2), foetal anomaly despite a normal karyotype or prior unexplained *in utero* foetal death at a gestational age ≥20 weeks.
**Score different according to the region of Asia (Central (0); Chinese, Southern, Maritime South East (1); Mainland South East (2)).

Reference standard

GDM diagnosis established according to the recommendations provided by the Canadian Diabetes Association in 2008 (50 g GCT in all women followed by 75 g OGTT if GCT between 7.8 and 10.2 mmol/L); test likely administered at 24–28 WG

Measures of test accuracy

Sensitivity, specificity, positive predictive value, negative predictive value and positive and negative likelihood ratios for predicting GDM along with their 95% CI were determined for different thresholds. Performance at the threshold maximising the Youden index (sensitivity/100 + specificity/100 - 1) was reported and AU(ROC) curves with their 95% CI were calculated. The prevalence of GDM for each risk score was compared graphically with the distribution obtained in the original studies.

GDM prevalence

A total of 381 participants developed GDM (5.3%). The prevalence of GDM increased with an increasing risk score for the four models, with data presented graphically in the publication.

Comparison of test accuracy between different screening methods

	Model	N participants/ n women with GDM	Sensitivity (%, 95% CI)	Specificity (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	AUC GDM (%, 95% CI)	Original cohort AUC GDM (%, 95% CI)*
Test Accuracy Outcomes	Naylor et al.	6160/324	72.2 (66.9–77.0)	55.1 (53.8–56.4)	8.2 (7.2–9.3)	97.3 (96.6–97.8)	0.668 (0.637–0.699)	0.729 (0.672–0.785)
	Caliskan et al.	5639/311	71.1 (65.6–76.0)	59.3 (58.0–60.6)	9.3 (8.1–10.5)	97.2 (96.6–97.8)	0.680 (0.649–0.712)	0.833 (0.735–0.930)
	van Leeuwen et al.	5302/280	60.4 (54.3–66.1)	80.7 (79.6–81.8)	14.9 (12.9–17.1)	97.3 (96.8–97.8)	0.756 (0.725–0.787)	0.770 (0.690–0.850)
	Teede et al.	4408/247	65.6 (59.3–71.4)	75.0 (73.7–76.3)	13.5 (11.6–15.6)	97.3 (96.7–97.9)	0.739 (0.701–0.776)	0.703 (0.646–0.759)

*AUC and CI calculated from data in the original articles.

- At the threshold maximising the Youden index, LR+ and LR- for the van Leeuwen et al. model were 3.13 (2.80–3.49) and 0.49 (0.43–0.57), respectively.

Study Reference **Theriault 2014**

- The models were also assessed for their power to predict the need for insulin therapy in GDM, as well as the ability to discriminate GDM cases requiring dietary intervention only and IGT cases from controls, with additional measures of test accuracy reported in the publication.

Authors' Conclusions	External validation in a large cohort of Caucasian women of four risk-prediction models based exclusively on clinical characteristics yielded a performance similar to those observed in the original studies, suggesting that the proposed models are generalisable. Modulation of the different variables to better reflect the population's characteristics, regarding for example ethnicity and body-mass index as well as inclusion of other factors such as physical activity could allow the creation of a more performing tool. The final model could serve to determine an a priori risk as part of an integrated screening strategy.

Abbreviations: WG, weeks of gestation.

Table 81: van Leeuwen 2010

van Leeuwen 2010

Design

Design

A model development study based on a prospective cohort study

Objective

To develop a multivariable logistic regression model which combines patient characteristics and medical history to predict the occurrence of GDM

Dates

NR

Country

Netherlands

Setting

A perinatal centre (University Medical Centre in Utrecht)

tion teristics

Patient recruitment and eligibility

Recruitment: All women who reported for prenatal care during the study period
Inclusion criteria: Singleton pregnancy
Exclusion criteria: Pregestational diabetes mellitus and women who were first seen after 20 weeks of gestation
Other: The original study was performed in two perinatal centres; however, for the development of the prediction model, only data from one centre was used (University Medical Centre in Utrecht)

Sample size

N screened/invited = NR
N eligible = NR
N enrolled = 995
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = 978 (had 50 g glucose challenge test)
N excluded from analysis = NR
N included in analysis = 995 women included in the original cohort study

Maternal demographics

NR

Maternal glycaemic characteristics NR

Index test/Comparator **Methods**
Model prediction based on patient characteristics and medical history. **van**

Leeuwen 2010 Reference standard

Random glucose testing and 50 g GCT performed in women once between the 24 and 28 weeks of gestation. If random plasma glucose ≥6.8 mmol/L or if 1-hour 50 g GCT ≥7.8 mmol/L, then a 2-hour 75 g OGTT was performed within 1 week. OGTT was performed in the morning after a 12-hour overnight fast and after 3 days of minimal 150–200 g carbohydrate diet. GDM was diagnosed if 2-hour venous plasma glucose ≥7.8 mmol/L, or if fasting blood glucose >7.0 mmol/L, according to WHO criteria.

Measures of test accuracy

ORs, 95% CIs and p-values were calculated from univariable analysis and from multivariable logistic regression analysis with a stepwise backwards selection procedure to construct the prediction model.

The discriminative performance of the model was assessed by ROC curve analysis and calculation of the AUC.

Sensitivity, specificity, positive predictive value, negative predictive value and likelihood ratios were reported.

Two thresholds of probability of GDM (2.0% and 4.0%) were evaluated.

GDM prevalence

A total of 24 of 995 women (2.4%) were diagnosed with GDM after correction for verification bias.

The mean AUC of the ROC curves from 10 multiple imputed datasets was 0.77 (95% CI 0.69–0.85), demonstrating a reasonable capacity to discriminate between women with and without GDM.

Comparison of test accuracy between different screening methods

Test Accuracy Outcomes	**Method of screening**	**Women to be tested with OGTT, n (%)**	**Sensitivity (%, 95% CI)**	**Specificity (%, 95% CI)**	**PPV (%, 95% CI)**	**NPV (%, 95% CI)**	**Likelihood ratio of a positive test result**	**Likelihood ratio of a negative test result**	**Women tested with OGTT (%)**	**OGTT to diagnose one case of GDM, n (N/n patients with GDM diagnosed)**
	Universal testing with OGTT	(100)	100	100	*NR*	*NR*	*NR*	*NR*	*NR*	(995/24)
	Diagnostic testing if the probability of GDM ≥2.0%	(43.0)	75.0 (55.4–88.0)	57.8 (57.3–58.1)	4.2 (3.1–4.9)	98.9 (98.1–99.5)	1.78 (1.30–2.10)	0.43 (0.21–0.78)	43.0	(428/18)
	Diagnostic testing if the probability of GDM ≥4.0%	(12.5)	45.8 (28.2–64.5)	88.4 (87.9–88.8)	8.9 (5.5–12.5)	98.5 (98.0–99.0)	3.94 (2.34–5.77)	0.61 (0.40–0.81)	12.5	(124/11)
	The accuracy measures of the current prediction model are additionally compared to two published scoring systems in the main body of the publication. Additionally, a nomogram to estimate the probability of GDM based on the presence of various different risk factors was presented graphically in the publication.									
Study Reference	**van Leeuwen 2010**									
Authors' Conclusions	An accurate clinical prediction model for pregnant women has been developed that can estimate the risk of GDM at booking based on patient characteristics. The use of a decision rule based on this prediction model could identify women at risk for GDM early in pregnancy, allowing for timely intervention to improve maternal and neonatal outcome.									

Question 3:

Data Extraction – Question 3: What is the most effective intervention for lowering glucose levels in screen-detected pregnant women with GDM and preventing adverse perinatal outcomes?

Table 82: Farrar 2016

Study Reference	**Farrar 2016 Chapter 6**
Study Reference	**Farrar 2016 Chapter 6**
Study Design	Design Systematic literature review Objective To evaluate the effects of oral anti-diabetic pharmacological therapies for treating women with GDM. Search dates 11th–12th September 2013, updated on 14th October 2014 Country Various Setting NR

Study eligibility

Inclusion (PICOS)

	Study Reference	Farrar 2016 Chapter 6
rence stics	**Farrar 2016**Population **Chapt**	**r 6** Pregnant women diagnosed with GDM or IGT using any threshold definition
	Intervention	Any one or more of: insulin, metformin, glibenclamide, dietary advice and diet modification with or without additional lifestyle modification (e.g. exercise) or monitoring
	Comparator	The comparison group could receive 'standard/routine obstetric care' (however defined by the trial) or any of the above treatments.
	Outcomes	Trials had to report incidence of adverse outcomes, for example RRs, ORs or mean differences (MDs) for outcomes compared across treatment groups for at least one of the following outcomes, which could be defined variously by the trials: • Gestational age at birth BW macrosomia (BW of ≥ 4 kg) • LGA (BW of > 90th centile) shoulder dystocia preterm birth < 37 weeks' gestation) • Neonatal hypoglycaemia • Admission to neonatal intensive care unit (NICU) • C-section (elective or emergency) • Pre-eclampsia • PIH induced labour • Instrumental birth (forceps or vacuum/ventouse) • Apgar score at 5 minutes • Negative treatment effects (e.g. gastrointestinal upset, well-being)
	Study design	RCTs (blinding of clinicians or researchers (to the intervention) or those assessing outcome data was not part of the inclusion criteria)

Exclusion (reasons given in excluded study list)

- Women with type 1 or type 2 diabetes
- Non-RCTs or quasi-randomised trials
- Postpartum intervention

Other

NR

Flow of Studies (PRISMA)

- Database results (Sept 2013 and Oct 2014): 6450 (2985 and 3555, respectively)
- Hand-searches/other sources: NR
- 3645 records after duplicates removed
- Title/abstracts reviewed: 3645
- Full-texts reviewed: 158
- Articles included in qualitative synthesis: 47

- Articles included in quantitative synthesis (meta-analysis): 45

Included study characteristics

Characteristic	Details
Design	RCTs
Sample sizes	Not summarised
Setting and timing	Not summarised
Participants	Not summarised
Diagnostic criteria for GDM	Variety of criteria used, including Carpenter and Coustan or NDDA, WHO, ADA, or local guidelines
Treatment targets	NR
Interventions and comparisons	23 trials compared drug treatments • Metformin vs. insulin: 10 trials • Glibenclamide vs. insulin: 8 trials • Glibenclamide vs. metformin: 2 trials • Metformin with glibenclamide vs. insulin: 1 trial • Glibenclamide with diet therapy vs. placebo with diet therapy: 1 trial 10 trials compared combinations of diet modification, glucose monitoring and insulin to routine obstetric care 5 trials compared different insulin formulations 9 trials compared different diets
Outcomes	Not summarised
Funding	NR
Conflicts of interest	NR

Definition of GDM

As defined in the individual trial

Searches Sources searched

- MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid SP)
- EMBASE (via Ovid SP)
- CENTRAL (via The Cochrane Library/Wiley Interscience)
- Reference searches of included journal articles and related systematic reviews

Screening and selection process

Quality assessment

Study Reference **Farrar 2016 Chapter 6**

Risk of bias was assessed using the Cochrane risk of bias tool. Each criterion was classed as being at low, high or unclear risk of bias. One reviewer performed the quality assessment, which was checked by a second. Methods for combining intervention evidence

Meta-analysis: The included trials were divided into categories according to the included treatments:

- Insulin vs. metformin
- Insulin vs. glibenclamide
- Metformin vs. glibenclamide
- Diet or dietary advice and or lifestyle vs. pharmacological (glibenclamide, metformin or insulin) treatment
- Diet or dietary advice and/or glucose monitoring and/or insulin use vs. routine antenatal care

The results of the trials comparing different types of insulin and different types of diet were not pooled because of their diversity and were reviewed narratively.

For dichotomous outcomes the RR for each outcome comparing each trial arm, with its 95% CI, was calculated from the numbers of women with the outcome. For continuous outcomes the MD between trial arms, with its 95% CI, was calculated from the mean and SD of the outcome. For each outcome, and within each of the four treatment categories listed above, RRs or mean differences were pooled in random-effects DerSimonian and Laird meta-analyses. Heterogeneity was assessed using Higgins I2-statistic. Subgroup analyses were performed to investigate differences across varying definitions of GDM.

Network meta-analysis

Network meta-analysis was used to combine information across multiple treatments simultaneously. Formally, analyses were conducted for each dichotomous outcome using a Bayesian approach, based on the models originally created by Lu and Ades, using the OpenBUGS software. Each model generated a comparison between treatments, expressed as an OR and a probability that each treatment was the best treatment to reduce the incidence of the outcome. Network meta-analysis was performed to compare insulin versus metformin versus glibenclamide.

Meta-analysis results: metformin vs insulin

Study Reference **Farrar 2016 Chapter 6**

Effectiveness of the **Intervention**

Outcome	Number of trials	I_2	Risk Ratio (95% CI)
Pregnancy outcomes			
Gestational age at birth			NR
Hypertensive disorders of pregnancy	3	0	0.54 (0.31–0.91)
Pre-eclampsia	4	0	0.74 (0.48–1.14)
Perinatal mortality	NR	NR	NR
Mode of birth	-	-	-
Induction of labour	4	52	0.84 (0.60–1.18)
Vaginal delivery	NR	NR	NR
Instrumental delivery	3	0	1.66 (1.37–2.01)
C-section (not specified if planned or emergency)	5	71	1.03 (0.66–1.62)
Maternal gestational weight gain	NR	NR	NR
Preterm delivery	4	50	1.37 (0.62–3.01)
Perineal trauma/tear	NR	NR	NR
Maternal outcomes	-	-	-
Maternal wellbeing	NR	NR	NR
Postpartum haemorrhage	NR	NR	NR
Method of infant feeding	NR	NR	NR
Postnatal depression	NR	NR	NR
Postnatal weight retention or return to prepregnancy weight	NR	NR	NR
Post-pregnancy T2DM	NR	NR	NR
Neonatal outcomes	-	-	-
Macrosomia	9	0	0.75 (0.57–0.98)
LGA	6	15	0.81 (0.62–1.05)
Adiposity (neonatal) – neonatal fat mass (g)	NR	NR	NR
Shoulder dystocia	3	0	0.99 (0.67–1.45)
Brachial plexus neuropathy	NR	NR	NR
Neonatal hypoglycaemia	7	0	0.71 (0.51–0.98)
Admission to NICU	8	60	0.79 (0.61–1.01)
Apgar score <7 at 5 minutes			3.06 (0.31–29.26)

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Study Reference **Farrar 2016 Chapter 6**

Meta-analysis results: glibenclamide vs. insulin

Outcome	Number of trials	I_2	Risk Ratio (95% CI)
Pregnancy outcomes			
Gestational age at birth			NR
Hypertensive disorders of pregnancy	NR	NR	NR
Pre-eclampsia	2	0	1.14 (0.60–2.18)
Perinatal mortality	NR	NR	NR
Mode of birth			-
Induction of labour	NR		NR
Vaginal delivery			NR
Instrumental delivery			NR
C-section (not specified if planned or emergency)	4	25	0.86 (0.66–1.12)
Maternal gestational weight gain			NR
Preterm delivery	1	0	0.50 (0.05–5.24)
Perineal trauma/tear			NR
Maternal outcomes			-
Maternal wellbeing			NR
Postpartum haemorrhage			NR
Method of infant feeding			NR
Postnatal weight retention or return to prepregnancy weight			NR
Post-pregnancy T2DM			NR
Neonatal outcomes			-
Macrosomia	4	29	2.66 (0.91–7.77)
LGA	5	59	2.44 (0.97–6.15)
Adiposity (neonatal) – neonatal fat mass (g)			NR

Study Reference **Farrar 2016 Chapter 6**

Shoulder dystocia			NR
Brachial plexus neuropathy			NR
Neonatal hypoglycaemia	4	0	1.60 (0.99–2.60)
Admission to NICU	2	0	0.95 (0.49–1.84)
Apgar score <7 at 5 minutes			NR

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Meta-analysis results: glibenclamide vs. insulin

Outcome	Number of trials	I_2	Risk Ratio (95% CI)
Pregnancy outcomes			
Gestational age at birth	NR	NR	NR
Hypertensive disorders of pregnancy	NR	NR	NR
Pre-eclampsia	2	0	1.14 (0.60–2.18)
Perinatal mortality	NR	NR	NR
Mode of birth			-
Induction of labour	NR	NR	NR
Vaginal delivery	NR	NR	NR
Instrumental delivery	NR	NR	NR
C-section (not specified if planned or emergency)	4	25	0.86 (0.66–1.12)
Maternal gestational weight gain	NR	NR	NR
Preterm delivery	1	0	0.50 (0.05–5.24)
Perineal trauma/tear	NR	NR	NR
Maternal outcomes			-
Maternal wellbeing	NR	NR	NR
Postpartum haemorrhage	NR	NR	NR
Method of infant feeding	NR	NR	NR
Postnatal weight retention or return to prepregnancy weight	NR	NR	NR
Post-pregnancy T2DM	NR	NR	NR
Neonatal outcomes			-

Study Reference **Farrar 2016 Chapter 6**

Macrosomia	4	29	2.66 (0.91–7.77)
LGA	5	59	2.44 (0.97–6.15)
Adiposity (neonatal) – neonatal fat mass (g)	NR	NR	NR
Shoulder dystocia	NR	NR	NR
Brachial plexus neuropathy	NR	NR	NR
Neonatal hypoglycaemia	4	0	1.60 (0.99–2.60)
Admission to NICU	2	0	0.95 (0.49–1.84)
Apgar score <7 at 5 minutes	NR	NR	NR

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Glibenclamide vs metformin

Dichotomous Outcome	Number of Trials	Risk ratio (95% CI)
LGA	1	2.29 (1.09–4.81)
Macrosomia	1	4.05 (0.46–35.42)
Admission to NICU admission	2	0.69 (0.29–1.66)
Neonatal hypoglycaemia	2	1.19 (0.57–2.48)
Shoulder dystocia	1	3.04 (0.13–73.44)
Continuous outcome	**Number of trials**	**Mean difference (95% CI)**
Gestational age at birth	2	0.11 (–0.65–0.86)
Birth weight	2	0.21 (–0.24–0.66)
Apgar score at 5 minutes	1	0.06 (–0.53–0.65)

Network meta-analysis comparing metformin, glibenclamide and insulin. First better – treatment listed first in the outcome column is superior; second better – treatment listed second in the outcome column is superior

Outcome	Number of Trials	Odds ratio (95% CI)
Metformin vs. insulin		
LGA	6	0.73 (0.42–1.27)
Macrosomia	9	0.64 (0.38–1.07)
Admission to NICU admission	8	0.62 (0.36–1.06)
Neonatal hypoglycaemia	7	0.71 (0.43–1.17)
Caesarean section	5	1.13 (0.55–2.30)
Pre-eclampsia	4	0.71 (0.34–1.47)

Study Reference **Farrar 2016 Chapter 6**

Glibenclamide vs. insulin		
LGA	5	2.37 (1.15–4.89)
Macrosomia	4	3.43 (1.32–8.91)
Admission to NICU	2	0.62 (0.25–1.59)
Neonatal hypoglycaemia	4	1.38 (0.74–2.60)
Caesarean section	4	0.70 (0.33–1.49)
Pre-eclampsia	2	1.19 (0.46–3.11)
Glibenclamide vs. metformin		
LGA	1	3.23 (1.41–7.40)
Macrosomia	1	5.38 (1.86–15.59)
Admission to NICU	2	1.01 (0.39–2.61)
Neonatal hypoglycaemia	2	1.95 (0.96–3.96)
Caesarean section	2	0.62 (0.26–1.51)
Pre-eclampsia	1	1.69 (0.56–5.11)

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Table 83. Estimated probability (%) of a treatment being the most effective in reducing the risk of a dichotomous outcome

Outcome	Treatment		
	Insulin	Metformin	Glibenclamide
LGA	17.2	82.7	0.1
Macrosomia	3.5	96.4	0
Admission to NICU	1.4	50.3	48.3
Neonatal hypoglycaemia	7.9	89.5	2.7
C-section	12.6	9.4	78
Pre-eclampsia	11.5	74.6	13.9

Study Reference **Farrar 2016 Chapter 6**

Meta-analysis results: Diet modification on dichotomous outcomes

Outcome	Number of Trials	I_2	Risk Ratio (95% CI)
Pregnancy outcomes			
Pre-eclampsia	5	33	0.58 (0.36–0.93)
Mode of birth	-	-	-
Induction of labour	4	66	1.12 (0.82–1.52)
Vaginal delivery	-	-	NR
Instrumental delivery	1	0	1.37 (0.20–9.27)
C-section (not specified if planned or emergency)	8	3	0.86 (0.77–0.95)
Preterm delivery	4	44	0.75 (0.46–1.21)
Maternal outcomes	-	-	-
NR	-	-	NR
Neonatal outcomes	-	-	-
Macrosomia	9	33	0.46 (0.35–0.60)
LGA	6	10	0.55 (0.44–0.69)
Shoulder dystocia	4	0	0.39 (0.23–0.69)
Neonatal hypoglycaemia	5	41	1.16 (0.79–1.69)
Admission to NICU	4	67	0.91 (0.62–1.34)
Apgar score <7 at 5 minutes	1	0	0.57 (0.21–1.52)

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Effect of different insulin preparation on dichotomous and continuous outcomes

Trial	Outcome	Risk ratio (95% CI)
Balaji	Caesarean section	1.00 (0.93–1.09)
	LGA	0.89 (0.39–2.04)
	Preterm birth	0.49 (0.04–5.36)
	Macrosomia	0.62 (0.16–2.37)
	Caesarean section	1.41 (0.57–3.49)

Study Reference **Farrar 2016 Chapter 6**

	LGA	1.33 (0.31–5.65)
	Macrosomia	1.50 (0.26–8.59)
	Neonatal hypoglycaemia	1.00 (0.31–3.24)
	C-section	1.01 (0.69–1.48)
	LGA	0.87 (0.59–1.27)
	Macrosomia	0.83 (0.50–1.40)
	Neonatal hypoglycaemia	0.12 (0.02–0.97)
	Pregnancy-induced hypertension	0.90 (0.41–1.98)
Trial	Outcome	Mean difference (95% CI)
	Apgar score at 5 minutes	0.01 (–0.11–0.13)
	Gestational age at birth	0.67 (0.33–1.01)
	Gestational age at birth	0.00 (–0.74–0.74)
	Apgar score at 5 minutes	0.30 (–0.30–0.90)
	Birthweight	–0.07 (–0.17–0.03)
	Gestational age at birth	0.00 (–0.41–0.41)
	Birthweight	0.10 (–0.08–0.28)
	Gestational age at birth	0.10 (–0.32–0.52)
	Birth weight	0.00 (–0.15–0.15)
	Gestational age at birth	0.30 (–0.12–0.72)

ent **Table 84. Risk of bias summary: review authors' judgements about each risk of bias item for each included**

Study Reference **Farrar 2016 Chapter 6**

Study	Included in a previous review	Random sequence generator (selection bias)	) Allocation concealment selection bias	) Blinding of participants and personnel performance bias	) Blinding of outcome assessment detection bias	Incomplete outcome data (attrition bias)	) Selective reporting reporting bias
Abbassi-Ghanavati 2014	-	Unclear	Low risk	Unclear	Low risk	High risk	Unclear

Anjalakshi 2007	-	Unclear	Unclear	Unclear	Unclear	Low risk	Unclear
Ardilouze 2014	-	Unclear	Unclear	Unclear	Unclear	Unclear	High risk
Asemi 2014	-	Low risk	Unclear	High risk	High risk	Low risk	Low risk
Balaji 2012	a	Unclear	Unclear	High risk	High risk	Low risk	Low risk
Bertini 2005	Alwan	Low risk	Low risk	High risk	High risk	Low risk	Low risk
Bevier 1999	Hartling	Unclear	Unclear	High risk	High risk	High risk	Low risk
Bo 2014	-	Low risk	Low risk	High risk	Low risk	Low risk	Low risk
Bonomo 2005		Unclear	Unclear	High risk	High risk	Low risk	Unclear
Bung 1991	-	Unclear	Unclear	High risk	High risk	High risk	Unclear
Cao 2012	-	Unclear	Unclear	High risk	Unclear	High risk	Low risk
Crowther 2005	Alwan, Falavigna, Hartling, Horvath	Low risk	Low risk	High risk	Low risk	Low risk	Low risk
Cypryk 2007	-	Unclear	High risk	Unclear	Unclear	Low risk	High risk
Deveer 2013	-	High risk	High risk	High risk	High risk	Low risk	Low risk
Di Cianni 2007	-	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Elnour 2008	a	Unclear	High risk	High risk	High risk	High risk	Low risk
Garner 1997	Falavigna, Hartling	Low risk	High risk	High risk	High risk	Low risk	Low risk
Hague 2003	Alwan	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Hassan 2012		High risk	High risk	Unclear	Unclear	Low risk	Low risk
Ijas 2010		Low risk	Low risk	High risk	High risk	Low risk	Low risk
Jovanovic 1999		Low risk	Unclear	High risk	High risk	Low risk	Low risk
Kjos 2001		Low risk	Unclear	High risk	Low risk	Low risk	Low risk
Lain 2009		Low risk	Low risk	Low risk	Low risk	High risk	Low risk
Landon 2009		Low risk	Low risk	High risk	Low risk	Low risk	Low risk
Langer 2000		Low risk	Unclear	Unclear	Unclear	Low risk	Low risk
Li 1987		High risk	Unclear	High risk	Unclear	Low risk	Low risk
Louie 2011		Low risk	Low risk	Low risk	Unclear	Low risk	High risk

Mesdaghinia 2013		Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	
Moore 2007		Low risk	Unclear	Unclear	Unclear	Low risk	Low risk	
Moore 2010		Low risk	Low risk	High risk	High risk	Low risk	Low risk	

Study MiG (Rowan 2018) Reference

Moreno-Castilla 2013		Unclear	Low risk	High risk	Unclear	Low risk	Low risk
Mukhopadhyay 2012		Low risk	Unclear	High risk	Unclear	Low risk	Low risk
Nachum 2012		Low risk	Low risk	Unclear	Low risk	Low risk	Low risk
Ogunyemi 2007		Low risk	Low risk	Unclear	Unclear	Low risk	Unclear
O'Sullivan 1966		Unclear	Unclear	High risk	High risk	Unclear	Unclear
Rae 2000		Unclear	Unclear	Low risk	Unclear	Low risk	High risk
Rowan 2008		Low risk	Unclear	High risk	High risk	Low risk	Low risk
Silvia 2012		Low risk	Unclear	High risk	High risk	Low risk	Low risk

Study Reference Farrar 2016 Chapter 6

Silva 2007		Unclear	Low risk	High risk	High risk	Low risk	Low risk
Spaulonci 2013		Low risk	Unclear	Unclear	Unclear	Low risk	Low risk
Tempe 2013		Unclear	Low risk	Unclear	Unclear	Low risk	Low risk
Tertti 2013		Unclear	Unclear	Unclear	Unclear	Low risk	Low risk
Thompson 1990		Low risk	Unclear	High risk	Unclear	High risk	Low risk
Yang 2003		Unclear	Unclear	High risk	Unclear	High risk	Unclear
Zinnat 2013		Unclear	Unclear	Unclear	Unclear	Low risk	Unclear

Treatment of GDM with diet and lifestyle and pharmacological interventions seems to reduce the risk of most reported perinatal adverse outcomes. Diet modification alone seems to reduce the risk of adverse outcomes even in women with glucose levels below those currently diagnostic of GDM. The provision of dietary advice for all pregnant women (irrespective of their glucose levels at OGTT) may be beneficial in terms of reducing the risk of adverse outcomes across the whole glucose spectrum.

Authors' Conclusions

Supplemental metformin in addition to diet and lifestyle modification (if required to normalise glucose levels) is as effective as insulin and therefore should be the first-line pharmacological treatment of choice, as it is at least as effective as insulin and may be preferred by women because it does not require injection, although it should be remembered that trials generally used insulin in the metformin group if hyperglycaemia was not 'well' controlled. A 'step-up' approach of first

providing dietary and lifestyle advice then adding supplementary metformin or insulin if glucose levels are not adequately controlled is a reasonable and effective approach to take.

Table 85: Metformin in Gestational Diabetes (MiG), Rowan 2018

Study	MiG (Rowan 2018) Reference
Study Design	Design RCT (longitudinal follow-up) Objective To compare body composition and markers of insulin sensitivity in offspring of women with GDM randomised to metformin (plus supplemental insulin as required) or insulin. Dates NR (follow-up at 7 and 9 years after birth of the offspring) Country Australia and New Zealand Setting Two sites in Adelaide and Auckland

Patient recruitment and eligibility

Recruitment: NR

Inclusion criteria: Women with GDM aged 18–45 years, at 22–33 weeks' gestation of a singleton pregnancy, who required pharmacotherapy based on capillary glucose levels.

Exclusion criteria: Diabetes pre-pregnancy, gestational hypertension, pre-eclampsia of foetal growth restriction at study entry, foetal congenital anomaly, maternal medical condition posing contra-indication to metformin.

Other: Longitudinal follow-up study includes offspring from the subset of trial participants who agreed to follow-up on an annual basis

Sample size

N screened/invited = NR
N eligible = NR
N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = 733 (completed the original trial)
N excluded from analysis = NR
N included in analysis = 208 (n=208 offspring) with follow-up data available

Maternal demographics **Population**

Characteristics

Characteristic	Total MiG cohort (n=733)	Subgroup seen at 7 years (Adelaide), (n=109)	Subgroup seen at 9 years (Auckland), (n=99)

Study **MiG (Rowan 2018) Reference**

		Metformin (n=58)	Insulin (n=51)	Metformin (n=45)	Insulin (n=54)
Age, years, mean ± SD	32.8 ± 5.3	33.6 ± 5.7	33.9 ± 4.7	34.12 ± 5.12	35.21 ± 4.72
Cardiometabolic health, mean ± SD					
BMI at booking (<20 weeks), kg/m^2	32.1 ± 7.9	31.3 ± 7.8	31.9 ± 8.3	31.1 ± 8.8	29.5 ± 6.4
BMI at enrolment, kg/m^2	34.9 ± 7.8	34.2 ± 7.1	34 ± 7.9	35.4 ± 11.3	32.0 ± 6.3
Weight, kg	NR	NR	NR	NR	NR
Ethnicity, n (%)					
European/Caucasian	343 (46.8)	52 (89.7)	43 (84.3)	25 (55.6)	21 (38.9)
Polynesian	156 (21.3)	0 (0)	0 (0)	6 (13.3)	7 (13.0)
Indian	93 (12.7)	0 (0)	4 (7.8)	7 (15.6)	16 (29.6)
Chinese and other Southeast Asian	86 (11.7)	4 (6.9)	2 (3.9)	6 (13.3)	7 (13.0)
Other or mixed	55 (7.5)	2 (3.4)	2 (3.9)	1 (2.2)	3 (5.6)
Medical history/risk factors, n (%)					
Chronic hypertension	58 (7.9)	7 (12.1)	5 (9.8)	7 (15.6)	5 (9.3)
Family history of diabetes (1st degree)	343 (46.8)	17 (29.3)	20 (39.2)	25 (55.6)	35 (64.8)

Study **MiG (Rowan 2018) Reference**

Smoking in pregnancy	121 (16.5)	7 (12.1)	1 (2.0)	5 (11.1)	4 (7.4)
Pre-pregnant alcohol use	NR	NR	NR	NR	NR
Obstetric history, n (%)					
Nulliparous	NR	NR	NR	NR	NR
Parous without GDM	NR	NR	NR	NR	NR
Parous with GDM	NR	NR	NR	NR	NR
Education level, n (%)					
Tertiary education	323 (44.1)	30 (51.7)	32 (62.7)	28 (62.2)	32 (59.3)

Definition of GDM NR

Methods

Duration of follow-up

7 years (Adelaide) and 9 years (Auckland) after birth of the offspring

Method of assigning treatment arm

Randomisation carried out online, using sequence blocks to stratify by site.

Arm 1 (n=58 in Adelaide, n=45 in Auckland)

Metformin (plus insulin if required)

Details NR

Arm 2 (n=51 in Adelaide, n=54 in Auckland)

Insulin only Details

NR

Outcomes

Primary endpoint

- Composite of neonatal morbidity including hypoglycaemia, respiratory distress, prematurity, phototherapy, birth trauma and low Apgar score

Secondary endpoints

- Maternal glycemia control
- Neonatal anthropometry
- Cord blood measures of adipoinsular axis
- Maternal hypertensive complications and postpartum glucose tolerance test results
- Acceptability of treatments

Study **MiG (Rowan 2018)**

Reference **Effectiveness of the Intervention**

Maternal, pregnancy and neonatal outcome data

Outcome	Subgroup seen at 7 years (Adelaide), (n=109)			Subgroup seen at 9 years (Auckland), (n=99)		
	Metformin (n=58)	Insulin (n=51)	p-value	Metformin (n=45)	Insulin (n=54)	p-value
Pregnancy outcomes						

Gestational age at birth, weeks, mean ± SD	38.4 ± 1.2	38.8 ± 1.0	0.05	38.4 ± 1.3	38.5 ± 1.2	0.75
Gestational age at birth <37 weeks, n (%)	6 (10.3)	2 (3.9)	0.28	5 (11.1)	6 (11.1)	1.00
Perinatal mortality	N/A	N/A	N/A	N/A	N/A	N/A
Mode of birth	--	--	0.44	--	--	0.83
Vaginal delivery	33 (56.9)	33 (64.7)	NR	30 (66.7)	34 (63.0)	NR
C-section	25 (43.1)	18 (35.3)	NR	15 (33.3)	20 (37.0)	NR
Maternal weight gain between enrolment and 36 weeks, kg, mean ± SD	1.0 ± 2.5	0.7 ± 2.4	0.59	0.4 ± 3.2	1.6 ± 2.8	0.07
Maternal outcomes, n (%)						
Gestational hypertension	1 (1.7)	0 (0)	1.00	5 (11.1)	3 (5.5)	0.46
Pre-eclampsia	3 (5.1)	2 (3.9)	1.00	2 (4.4)	0 (0)	0.20
Maternal wellbeing	NR	NR	NR	NR	NR	NR
Postpartum haemorrhage	NR	NR	NR	NR	NR	NR
Method of infant feeding						
Breast feeding	32 (55.1)	25 (49.0)	0.30	25 (55.6)	30 (56.6)	0.59
Formula feeding	17 (29.3)	13 (25.5)		5 (11.1)	10 (18.9)	
Both breast and formula feeding	5 (8.6)	13 (25.5)		14 (31.1)	13 (24.5)	
Not seen	4 (6.8)	0 (0)		1 (2.2)	1 (1.9)	
Post-pregnancy T2DM (selfreported)	NR	NR	NR	19 (42.2)	22 (40.7)	1.00
Neonatal outcomes						

Study Reference **MiG (Rowan 2018)**

Other outcomes reported

Supplementary insulin prescribed, glycaemic control from randomisation to delivery (mean fasting capillary glucose, mean postprandial glucose, mean glucose, HbA1c at 36 weeks), birth weight customised centile, crown-heel length, crown-rump length, head circumference, chest circumference, abdominal circumference, mid-upper arm circumference, triceps skinfold thickness, subcapsular skinfold thickness, ponderal index.

Authors' Conclusions	No relevant conclusions (publication focused on outcomes at 7 and 9 years).						
	Macrosomia	NR	NR	NR	NR	NR	NR
	Birth weight, g, mean ± SD	3,481 ± 565	3,324 ± 431	0.10	3,284 ± 563	3,238 ± 542	0.69
	Birth weight >90th percentile, n (%)	12 (20.7)	3 (5.9)	0.029	5 (11.1)	6 (11.1)	1.00
	Birth weight <10th percentile, n (%)	5 (8.6)	4 (7.8)	1.0	5 (11.1)	6 (11.1)	1.00
	Birth injury	NR	NR	NR	NR	NR	NR
	Shoulder dystocia	NR	NR	NR	NR	NR	NR
	Brachial plexus neuropathy	NR	NR	NR	NR	NR	NR
	Neonatal hypoglycaemia	NR	NR	NR	NR	NR	NR
	Admission to NICU	NR	NR	NR	NR	NR	NR

Abbreviations: BMI: Body Mass Index; GDM: gestational diabetes mellitus; HbA1c: glycated haemoglobin A1c; LGA: large for gestational age; MiG: Metformin in Gestational Diabetes; N/A: not applicable; NICU: newborn intensive care unit; NR: not reported; RCT: randomised controlled trial; SD: standard deviation; SGA: small for gestational age; T2DM: type 2 diabetes mellitus.

Table 86: Hernandez 2016

Study **MiG (Rowan 2018)** **Reference** **Study Reference** **Hernandez 2016**

Study Design	Design RCT (pilot) Objective To test the hypothesis that compared with a conventional low-carbohydrate/higher-fat diet, consumption of a higher-complex carbohydrate/lower-fat diet would improve maternal insulin resistance, adipose tissue lipolysis and infant adiposity. Dates NR Country USA Setting NR

Population Characteristics

Patient recruitment and eligibility

Recruitment: NR

Inclusion criteria: Women diagnosed with GDM at 24–28 weeks using Carpenter-Coustan criteria, aged 20–36 years, with BMI 26–39 kg/m^2 and fasting blood glucose <110 mg/dL, treated with diet alone (receiving no other treatment) **Exclusion criteria:** Any comorbidities. **Other:** NR

Sample size

N screened/invited = NR

rence **Hernandez 2016**

N eligible = NR
N enrolled = 12
N excluded (with reason) = NR

N lost to follow-up = 0
N completed = 12
N excluded from analysis = 0 N included in analysis = 12

Maternal demographics at 31 weeks' gestation

Characteristic	CHOICE diet (n=6)	LC/CONV diet (n=6)
Age, years, mean ± SEM	30 ± 1	28 ± 2
Cardiometabolic health		
Pre-pregnant BMI, kg/m^2	NR	NR
BMI at study entry, kg/m^2	34.3 ± 1.6	33.4 ± 1.4
Weight at study entry, kg	91.2 ± 5.8	86.5 ± 5.1
Ethnicity, n		
Caucasian	4	2
Asian	2	1
Hispanic	0	3
Medical history/risk factors, n (%)		
Hypertension	NR	NR
Diabetes	NR	NR
Pre-pregnant smoking	NR	NR
Pre-pregnant alcohol use	NR	NR
Obstetric history, n		
Gravida	3	2
Parous	1	1
Parous without GDM	NR	NR
Parous with GDM	NR	NR
Education level	NR	NR

Definition of GDM

Carpenter-Coustan criteria

Study Reference	Hernandez 2016

Methods

Duration of follow-up

From gestational weeks 30–31 to 2 weeks postnatal

Method of assigning treatment arm NR

Arm 1 (n=6)

Higher-complex carbohydrate/lower-fat (CHOICE) diet

Composed of 60% carbohydrate/25% fat/15% protein. Menus were prepared by the research centre nutrition serviced and picked up by participants every 72 hours.

Arm 2 (n=6)

Low-carbohydrate/higher-fat (LC/CONV) diet

Composed of 40% carbohydrate/45% fat/15% protein, and matched with CHOICE diet for fat, simple sugars and fibre content. Menus were prepared by the research centre nutrition serviced and picked up by participants every 72 hours.

Outcomes

Endpoints (not specified as primary or secondary)

- Maternal insulin resistance, assessed as a measure of insulin suppression of adipose tissue lipolysis (where higher suppression indicates lower resistance), at 37 weeks' gestation
- Maternal proinflammatory gene expression at 37 weeks' gestation
- Maternal fasting free fatty acids at 37 weeks' gestation
- Infant body composition (weight, length and adiposity) at 2 weeks after birth
- Maternal fasting and 2-hour postprandial glucose throughout pregnancy (self-assessed)

Outcome	CHOICE diet (n=6)	LC/CONV diet (n=6)	p-value
Pregnancy outcomes			
Gestational age at birth, weeks, mean ± SEM	40.5 ± 0.5	39.2 ± 0.4	NR
Perinatal mortality	NR	NR	NR
Mode of birth, n			
Induction of labour	NR	NR	NR
Vaginal delivery	NR	NR	NR

Hernandez 2016

Preterm delivery	NR	NR	NR
Instrumental delivery	NR	NR	NR
C-section	0	2	NR
Maternal gestational weight gain (while on study), kg, mean ± SEM	2.3 ± 1.2	1.7 ± 1.6	NR

Study Reference

Effectiveness of Maternal, pregnancy and neonatal outcomes **the Intervention**

Maternal outcomes, n

	NR	NR	NR
Maternal wellbeing			
Postpartum haemorrhage	NR	NR	NR
Method of infant feeding	NR	NR	NR
Post-pregnancy T2DM	NR	NR	NR
Neonatal outcomes, mean ± SEM			
Weight, g	--	--	--
At delivery	3,273.0 ± 104.0	3,421.0 ± 186.3	NR
2 weeks postnatal	3,452 ± 113	3,683 ± 292	NR
Adiposity at 2 weeks postnatal, g	392 ± 43	510 ± 124	NR
Body fat at 2 weeks postnatal, %	10.1 ± 1.4	12.6 ± 2.0	NR
Macrosomia	NR	NR	NR

	LGA	NR	NR	NR
	Birth injury	NR	NR	NR
	Shoulder dystocia	NR	NR	NR
	Brachial plexus neuropathy	NR	NR	NR
	Neonatal hypoglycaemia	NR	NR	NR
	Admission to NICU	NR	NR	NR

Study Reference	**Hernandez 2016**
	Other outcomes reported BMI at delivery, weight at delivery, infant length at delivery, infant fat-free mass at 2 weeks, maternal lipid concentrations, maternal glucose concentration, maternal insulin concentration, maternal HOMA-IR, maternal pro-inflammatory gene expression.
Authors' Conclusions	The data suggest that the CHOICE diet, which is characterised by including more nutritious complex carbohydrates, may improve adipose tissue insulin resistance, resulting in a lower fasting blood glucose and free fatty acid exposure to the foetus. Therefore, adipose insulin action may be an important target for diet therapy in GDM.

Abbreviations: BMI: body Mass Index; CHOICE: higher-complex carbohydrate/lower-fat; GDM: gestational diabetes mellitus; HOMA-IR: homeostatic model assessment for insulin resistance; LC/CONV: low-carbohydrate/higher-fat; LGA: large for gestational age; NICU: newborn intensive care unit; NR: not reported; RCT: randomised controlled trial; SEM: standard error of the mean; T2DM: type 2 diabetes mellitus; USA: United States of America.

Table 87: Insulin Daonil (INDAO) trial, Senat 2018

Study Reference	**INDAO (Senat 2018)**
Study Design	Design RCT Objective To compare oral glyburide vs subcutaneous insulin in prevention of perinatal complications in newborns of women with GDM. Dates May 2012–November 2016 Country
Study Reference	**INDAO (Senat 2018)**

France

Setting

13 tertiary care hospitals

stics

Patient recruitment and eligibility

Recruitment: NR

Inclusion criteria: Women with a singleton pregnancy diagnosed as having GDM between 24 and 34 weeks of gestation, who failed to meet glycaemic goals (fasting: <95 mg/dL; 2-hour postprandial: <120 mg/dL) after 10 days of individual nutrition education by a dietitian.

Exclusion criteria: Pre-gestational diabetes, fasting blood glucose >126 mg/dL, glucose screening test performed before 24 weeks of gestation, multiple pregnancy, chronic hypertension, pre-eclampsia or known liver or renal disease.

Other: NR

Sample size

N screened/invited = NR

N eligible = NR

N enrolled = 914

N excluded (with reason) = 18 who did not meet the inclusion criteria, 6 who did not have data for the primary outcome or refused the treatment. 81 (18%) randomised to glyburide switched to insulin.

N lost to follow-up = NR

N completed = 809 (glyburide arm: 367; insulin arm: 442)

N excluded from analysis = 0

N included in analysis = 809 (glyburide arm: 367; insulin arm: 442) (per-protocol)

Maternal demographics (per-protocol)

Characteristic	Glyburide (n=367)	Insulin (n=442)
Age, years, mean (SD)	32.5 (5.1)	32.6 (5.3)
Cardiometabolic health, mean (SD)		
Pre-pregnancy BMI, kg/m^2	27.3 (5.5)	27.8 (5.8)
BMI at diagnosis, kg/m^2	30.7 (5.1)	31.1 (5.4)
Weight, kg	NR	NR
Geographical origin, n (%)		
Europe	146 (41.3)	184 (43.2)
North Africa	124 (34.9)	136 (31.9)
Sub-Saharan Africa	35 (9.9)	50 (11.7)

rence	INDAO (Senat 2018)		
	Asia	19 (5.4)	21 (4.9)
	Other	31 (8.7)	35 (8.2)
	Medical history/risk factors, n (%)		
	Hypertension	NR *(exclusion criterion)*	NR *(exclusion criterion)*
	Diabetes	NR *(exclusion criterion)*	NR *(exclusion criterion)*
	Pre-pregnant smoking	NR	NR
	Pre-pregnant alcohol use	NR	NR
	Obstetric history, n (%)		
	Parous	216 (58.9)	293 (66.3)
	Parous with GDM	73 (20.0)	88 (19.9)
	Education level	NR	NR

Definition of GDM

One or more abnormal blood glucose values (>92 mg/dL, >180 mg/dL or >153 mg/dL for fasting, 1-hour postprandial, or 2-hour postprandial blood glucose concentration, respectively) on a 75-g oral glucose tolerance test.

Methods	Duration of follow-up

Study Reference	**INDAO (Senat 2018)**

	NR Method of assigning treatment arm Eligible women were randomly assigned to one of the two treatment arms in a 1:1 ratio, using an independent, centralised, computer-generated randomisation sequence according to a permuted-block method with block sizes randomly chosen from 2 to 8, stratified by centre. Clinicians and participants had no access to the list but could not be blinded to group allocation after randomisation. Arm 1 (n=367) **Glyburide** Starting dosage for therapy was 2.5 mg orally once per day and could be increased if necessary 4 days later by 2.5 mg and thereafter by 5 mg every 4 days in 2 doses, morning and evening, up to 20 mg/d. If the maximum tolerated dosage was reached without achieving the desired glucose values of <95 mg/dL for fasting measurements and <120 mg/dL for 2-hour postprandial measurements, treatment was switched to insulin. Arm 2 (n=442) **Insulin** The starting dosage for rapid analogues was 4 IU given subcutaneously before meals, 1 to 3 times per day as necessary and increased by 2 IU every 2 days according to the postprandial blood glucose value. If necessary, the starting dosage for basal or intermediate insulin was 4 IU to 8 IU given subcutaneously at bedtime and increased by 2 IU every 2 days according to the morning fasting blood glucose value. Women were taught to self-adjust their insulin doses to reach and maintain glycaemic goals throughout pregnancy. Outcomes **Primary endpoint** • Composite criterion of perinatal complications associated with GDM, including macrosomia (birth weight >4,000 g or >90th percentile for gestational age according to French curves), neonatal hypoglycaemia (blood glucose <36 mg/dL after 2 hours of life) and hyperbilirubinemia (need for phototherapy without another cause of jaundice) **Secondary** **endpoints** • Neonatal perinatal death • Admission to NICU • Admission to neonatal ward
	• Respiratory distress syndrome • Birth injury • Ponderal index ([birth weight, g / (height, cm)^3] x 100) • Neonatal pH <7 • Lactate levels • Base excess (not recorded)
Study Reference	**INDAO (Senat 2018)**
	• Maternal glycaemic control during pregnancy (percentage of measurements ≥95 mg/dL [fasting] and ≥120 mg/dL [2 hours postprandial])

- Maternal hypoglycaemia (blood glucose >60 mg/dL) and/or symptomatic episode of hypoglycaemia with clinical symptoms of severity (confusion, poor coordination, double vision, convulsion or inability to self-treat symptoms) • Premature delivery
- Mode of delivery

- Perineal trauma
- Percentage switch from glyburide to insulin
- Maternal satisfaction (preferred treatment for future pregnancy)
- Number of prenatal visits
- Number of diabetologist visits

- Hospitalisation days during pregnancy

Outcome	Glyburide (n=367)	Insulin (n=442)	p-value
Pregnancy outcomes			
Gestational age at birth	NR	NR	NR
Perinatal mortality	0	2	NR
Mode of birth, n (%)	--	--	0.08
Spontaneous vaginal	(55.9)	(56.8)	NR
Assisted vaginal	(17.2)	(15.2)	NR
Emergency C-section	(17.2)	(13.1)	NR
Elective C-section	(9.8)	(14.9)	NR
Maternal gestational weight gain	NR	NR	NR
Preterm delivery, n (%)	(6.8)	(4.1)	0.09
Maternal outcomes, n (%)			
Maternal wellbeing	NR	NR	NR
Postpartum haemorrhage	NR	NR	NR
Perineal trauma	(0.8)	(0.2)	0.27
Method of infant feeding	NR	NR	NR
Post-pregnancy T2DM	NR	NR	NR

INDAO (Senat 2018)

Study Reference

Effectiveness of Intervention

Maternal, neonatal and pregnancy outcomes (per-protocol) **the**

Neonatal outcomes,

Birth weight, g, mean (SD)	3,341 (513)	3,331 (476)	0.77
Macrosomia (including LGA, see definition above), n (%)	59 (16.2)	65 (14.8)	0.59
Birth weight >4,000 g, n (%)	33 (9.3)	28 (6.6)	0.16
Hypoglycaemia, n (%)	45 (12.2)	32 (7.2)	0.02
Hyperbilirubinemia, n (%)	14 (3.8)	14 (3.1)	0.61
Birth injury, n (%)	6 (1.5)	9 (1.9)	0.66
Shoulder dystocia	1	2	NR
Bone fracture	1	6	NR
Nerve palsy	1	0	NR
Other[a]	3	1	NR
Neonatal hypoglycaemia	NR	NR	NR
Severe respiratory distress syndrome, n (%)	8 (1.9)	11 (2.2)	
Admission to NICU before 48h of life, n (%)	10 (2.3)	11 (2.4)	
	27 (7.9)	34 (8.2)	0.75
Admission to neonatal ward, n (%)	3 (0.8)	11 (2.5)	0.87
Apgar score ≤7 at 5 min, n (%)			0.86
			0.08

	[a] Glyburide group: facial haematoma, serosanguine hump, scalp wound; insulin group: serosanguine hump. Note: ITT analysis also available but not extracted here Other outcomes available: Infant pH<7, infant lactates concentration, glycaemic control during pregnancy (fasting, postprandial), maternal hypoglycaemia during pregnancy, preferred treatment for a future pregnancy.
Study Reference	**INDAO (Senat 2018)**
Authors' Conclusions	This study of women with gestational diabetes failed to show that use of glyburide compared with subcutaneous insulin does not result in a greater frequency of perinatal complications. These findings do not justify the use of glyburide as a first-line treatment.

Abbreviations: BMI: Body Mass Index; GDM: gestational diabetes mellitus; ITT: intention-to-treat; IU: International Unit; LGA: large for gestational age; NICU: newborn intensive care unit; NR: not reported; RCT: randomised controlled trial; SD: standard deviation; T2DM: type 2 diabetes mellitus.

Table 88: Trout 2016

Study Reference	**Trout 2016**
Study Design	Design RCT Objective To examine the effects of a maternal carbohydrate-restricted diet on maternal and infant outcomes in GDM. Dates NR Country

rence	**Trout 2016**
	USA Setting Two sites, including an urban teaching hospital (Site A) and a suburban community hospital (Site B)

Patient recruitment and eligibility

Recruitment: Participants were informed of the study at diagnosis of GDM; those interested in participating were screened for eligibility via a questionnaire administered by a study team member.

Inclusion criteria: Women aged 18–45 years, diagnosed with GDM by Carpenter-Coustan criteria at ≤35 weeks of gestation, whose GDM was controlled by diet alone or diet plus oral medication (e.g. glyburide).

Exclusion criteria: Multifetal pregnancy, pregestational diabetes or requiring insulin at the time of enrolment, any other significant medical or psychiatric comorbidities (e.g. cardiovascular disease or pre-existing hypertension), smoking, use of alcohol or illicit drugs. **Other:** NR

Sample size

N screened/invited = NR
N eligible = NR
N enrolled = 68
N excluded (with reason) = NR

N lost to follow-up = NR
N completed = NR
N excluded from analysis = 0
N included in analysis = 68 (ITT)

stics

Maternal demographics

Characteristic	Lower-carbohydrate diet (n=37)	Usual pregnancy diet (n=31)
Age at delivery, years, mean ± SD	30.09 ± 6.15	29.63 ± 5.19
Cardiometabolic health, mean ± SD		
Pre-pregnant BMI, kg/m^2	NR	NR
BMI, kg/m^2	33.84 ± 8.84	31.80 ± 8.68
Weight, kg	NR	NR
Ethnicity, n (%)	NR	NR
Medical history/risk factors, n (%)		
Pre-existing hypertension	NR *(exclusion criterion)*	NR *(exclusion criterion)*
Diabetes	NR *(exclusion criterion)*	NR *(exclusion criterion)*
Pre-pregnant smoking	NR	NR
Pre-pregnant alcohol use	NR	NR
Obstetric history, n (%)	NR	NR

Study Reference	**Trout 2016**

Education level	NR	NR

Definition of GDM

Women were screened with a 50-g OGTT at 24–28 weeks' gestation without regard to time of day or interval since their last meal. Women whose 1hour screening test plasma glucose result was ≥135 mg/dL underwent a 100-g OGTT and were diagnosed with GDM based on the Carpenter-Coustan criteria.

Methods

Duration of follow-up

NR

Method of assigning treatment arm

Randomised; details NR

Arm 1 (n=37)

Lower-carbohydrate diet

Participants received a total maximum daily carbohydrate gram count that was set after consultation with a certified diabetes educator in relation to each participant's total daily caloric intake. Participants were instructed on minimum and maximum recommended carbohydrate levels (35–40% of total calories, respectively) and received detailed instructions on carbohydrate gram counting, including the use of measuring cups, a calibrated gram scale and a list of carbohydrate grams for portions of common foods. Adherence was ensured through review of food logs.

Arm 2 (n=31)

Usual pregnancy diet

Participants underwent identical procedures in all aspects of the study, including carbohydrate counting and recording food intake, but had a carbohydrate intake level set at 50–55% of total calories.

Outcomes

Endpoints (not specified as primary or secondary)

- Maternal blood glucose control
- Maternal weight gain
- Composite maternal medical complications (including gestational hypertension, pre-eclampsia, polyhydramnios, post-partum haemorrhage and urinary tract infection)
- Incidence of maternal medical procedures
- Infant birth weight
- Incidence of macrosomia
- Composite infant adverse perinatal events (including macrosomia [birth weight ≥4,000 g], shoulder dystocia, respiratory distress syndrome, hypoglycaemia and admission to the NICU.

Effectiveness of	Maternal, neonatal and pregnancy outcomes (ITT)			
	come	Usual pregnancy diet (n=31)	p-value	**the Intervention Pregnancy outcomes**

ference	Trout 2016			
	Gestational age at birth, weeks, mean ± SD	37.78 ± 1.66	37.76 ± 1.74	0.96
	Perinatal mortality, %	0	0	NR
	Mode of birth, %			
	Induction of labour	35.3	34.4	0.94
	Vaginal delivery	NR	NR	NR
	Instrumental delivery	NR	NR	NR
	C-section	NR	NR	NR
	Primary C-section	29.4	40.6	0.34
	Emergency C-section	NR	NR	NR
	Planned C-section	NR	NR	NR
	Maternal gestational weight gain (total), lb, mean ± SD	27.24 ± 16.02	25.68 ± 17.08	0.71
	Maternal gestational weight gain (from study entry), lb, mean ± SD	4.75 ± 6.20	4.41 ± 6.24	0.85
	Preterm delivery	NR	NR	NR
	Maternal outcomes			
	Maternal wellbeing	NR	NR	NR
	Postpartum haemorrhage	NR	NR	NR
	Method of infant feeding	NR	NR	NR
	Post-pregnancy T2DM	NR	NR	NR
	Composite maternal complications[a], %			
	None	52.0	50.0	0.81
	≥1	47.1	50.0	
	Neonatal outcomes			
	Macrosomia, %	11.8	12.5	0.93
	LGA	NR	NR	NR
	Birth weight, g, mean ± SD	3,409.53 ± 527.91	3,377.28 ± 589.91	0.81
	Birth injury, %	NR	NR	NR
	Shoulder dystocia	2.9	0	0.25
	Bone fracture	0	0	NR
	Nerve palsy	0	0	NR
	Brachial plexus neuropathy	NR	NR	NR
	Neonatal hypoglycaemia, %	9.7	26.9	0.09
	Admission to NICU	20.6	12.5	0.38
	Composite infant complications, %	--	--	0.92
	None	67.6	68.8	NR
	≥1	32.4	31.1	NR

[b]

[a] Composite outcome included gestational hypertension, pre-eclampsia, polyhydramnios, post-partum haemorrhage, and urinary tract infection. Composite outcome included macrosomia (birth weight >4000 g), shoulder dystocia, respiratory distress syndrome, hypoglycaemia and admission to the NICU.

Study Reference **Trout 2016**

Other outcomes available:

Study Reference	Trout 2016
	Maternal 2-hour postpartum blood glucose concentration, fasting blood glucose concentration, maternal need for insulin therapy, oral medication use before and after enrolment, infant head circumference, infant abdominal girth.
Authors' Conclusions	No differences found between the lower-carbohydrate and usual-care diets in terms of blood glucose or maternal-infant outcomes; however, the results may not have been significant due to a type II resulting from too small a sample size.

Abbreviations: BMI: Body Mass Index; GDM: gestational diabetes mellitus; ITT: intention-to-treat; LGA: large for gestational age; NICU: newborn intensive care unit; NR: not reported; OGTT: oral glucose tolerance test; RCT: randomised controlled trial; SD: standard deviation; T2DM: type 2 diabetes mellitus; USA: United States of America.

Table 89: Pellonpera 2016, Huhtala 2018

Study Reference	Pellonpera 2016, Huhtala 2018
Study Design	Design An open-label randomized clinical trial Objective To investigate the possible effect of metformin treatment in women with GDM on weight gain and glucose tolerance compared with insulin or diet-only treatments at 6–8 weeks and 1 year postpartum. To compare concentrations of maternal amino acids and lactate in women diagnosed with GDM treated with metformin and insulin. Dates Between June 2006 and December 2010. Country Finland Setting Turuku University Hospital
	Patient recruitment and eligibility **Recruitment:** 228 women were recruited before change in GDM diagnostic criteria (December 2008) and 91 women after the change in the GDM diagnostic criteria (December 2008). **Inclusion criteria:** Women with singleton pregnancies who had at least two pathologic values in a 2-h 75-g OGTT and needed medication for the treatment of GDM (not achieving fasting glucose <5.5 and postprandial glucose <7.8 mmol/L during diet).

Population Characteristics Other: NR

Exclusion criteria: Cardiac or renal insufficiency, liver disease, metformin use within 3 months preceding pregnancy, or during pregnancy before the OGTT.

Study Reference **Pellonpera 2016, Huhtala 2018**

Sample size (Pellonpera 2016)
N screened/invited = NR
N eligible = NR

N enrolled = 221
N excluded (with reason) before randomisation = 4 (no reason provided)

N excluded (with reason) before post-partum follow-up = 12 (did not attend postpartum appointments)
N completed = 205 completed postpartum follow-up (N=104 metformin, N=101 insulin)

N excluded from analysis = Variable (as shown in tables)
N included in analysis = Variable (as shown in tables)

Sample size (Huhtala 2018)
N screened/invited = NR
N eligible = NR
N enrolled = 319
N excluded (with reason) or lost to follow up = 107 (no reason given)

N completed = 217 (N=110 metformin, N = 107 insulin)
N excluded from analysis = Variable depending on sample availability (as shown in tables) N included in analysis = Variable depending on sample availability (as shown in tables)

Maternal demographics

Characteristic	Metformin (N=110)		Insulin (N=107)		p-value
	Result	n[a]	Result	n[a]	
Age, years ± SD	31.9 ± 4.99	104	32.0 ± 5.49	101	NR[b]
Cardiometabolic health					
Pre-pregnant BMI, kg/m^2	29.5 ± 5.91	103–109	28.9 ± 4.71	101–107	0.41
BMI, kg/m^2	NR	NR	NR	NR	NR
Weight at first antenatal visit, kg ± SD	81.4 ± 16.5	98	80.4 ± 15.2	99	NR[b]
OGTT 0 h, mmol/L[c]	5.48 ± 0.53	104	5.53 ± 0.41	101	NR[b]
OGTT 1 h, mmol/L[c]	11.14 ± 1.49	104	11.02 ± 1.29	101	NR[b]
OGTT 2 h, mmol/L[c]	8.24 ± 1.76	102	7.78 ± 1.76	100	NR[b]
HbA1c, %[c]	5.47 ± 0.33	103	5.50 ± 0.33	100	NR[b]

Study Reference **Pellonpera 2016, Huhtala 2018**

Ethnicity, n (%)					
White/Caucasian	99% of all patients included in the trial				NR
Black	NR				NR
South Asian	NR				NR
East Asian	NR				NR
Mixed	NR				NR
Medical history/risk factors, n (%)					
Hypertension	NR	NR	NR	NR	NR
Diabetes	NR	NR	NR	NR	NR
Smoking	9 (8.6)	103–109	17 (16.0)	101–107	0.099
Pre-pregnant alcohol use	NR	NR	NR	NR	NR
Family history of DM, first-degree relative	32 (35.2)	91	30 (33.0)	91	NR[b]
Family history DM, first- or seconddegree relative	64 (70.3)	91	63 (69.2)	91	NR[b]
Obstetric history, n (%)					
Nulliparous	NR	NR	NR	NR	NR
Parous without GDM	NR	NR	NR	NR	NR
Parous with GDM	NR	NR	NR	NR	NR
Primipara	40 (38.5)	104	44 (43.6)	101	NR[b]

[a]Data available for n number of participants.
[b]No comparisons between metformin and insulin only. [c]Measured at the time of GDM diagnosis.

Definition of GDM

Up to December 2008, 2 or more values on OGTT exceeding: ≥4.8 mmol/L (fasting), 10.0 mmol/L (1 h) and ≥8.7 mmol/L (2 h). After December 2008, 2 or more values on OGTT exceeding: ≥5.3 mmol/L (fasting), 10.0 mmol/L (1 h) and ≥8.6 mmol/L (2 h).

Study Reference	Pellonpera 2016, Huhtala 2018
Methods	Duration of follow-up Up to 1 year postpartum Method of assigning treatment arm The women who had insufficient glycaemic control with diet treatment were randomized between 22 and 34 weeks of gestation to receive eith metformin or insulin medication. Women achieving sufficient glycaemic control (fasting glucose <5.5 and postprandial glucose <7.8 mmol/L) v only were not randomized for medication. Arm 1 (n=110) **Metformin** Metformin was started at 500 mg daily and increased up to 2000 mg if needed (median 1500 mg). Medication was initiated at mean 30 gw. If values of <5.5 mmol/L for overnight fasting plasma glucose and <7.8 mmol/L for 60-min postprandial plasma glucose were not met in the metformintreated women, additional insulin medication was started. In this group, 23 women needed additional insulin, of whom 21 were inclu follow-up analysis. Arm 2 (n=107) **Insulin** Insulin treatment was accomplished using NPH insulin and/or rapid-acting insulin lispro or insulin aspart. Medication was initiated at mean 30 Arm 3 (n=128) **Diet and lifestyle modifications** Diet and lifestyle modification advice was provided by public health nurses during normal antepartum visits in local maternity clinics. In additio women received dietary counselling in the university hospital. Outcomes **Endpoints (not specified as primary or secondary)**
	• Maternal weight gain, defined as the difference in weight between weight measured at the first antenatal visit at 8–9 weeks of gestatio measured at 6–8 weeks and 1 year after delivery, as measured by an electric scale. • Postpartum glucose tolerance, defined by 2-h OGTT and HbA1c levels, measured with the same assay as during gestation. The per women having IFG, IGT or DM were calculated using the criteria of the WHO, where IFG is defined as fasting plasma glucose 6.1–6 IGT as OGTT 2-h value plasma glucose 7.8–11.0 mmol/L and DM is defined as fasting plasma glucose ≥7.0 mmol/L or OGTT 2-h va glucose >11.0 mmol/L or HbA1c ≥6.5%. • Breastfeeding, defined by obtaining information from participants. • Concentrations of maternal amino acids and lactate. • Gestational weight gain, defined as last measured weight at maternity clinic minus self-reported weight before pregnancy.

Study Reference **Pellonpera 2016, Huhtala 2018**

- Gestation length
- The incidence of caesarean section
- Birth weight, defined in absolute value in grams and also calculated in SD units of deviation from Finnish general population mean adjusted for gestation length.
- Macrosomia, defined as birth weight >2.0 SD and/or >4500 g.
- SGA, defined as birth weight <–2.0 SD.
- Indicators of low and high birth weight, calculated as prevalence of birth weight >90th and <10th percentile of population adjusted for gestation length.
- Neonate admission to NICU
- Neonatal intravenous glucose given for any indication
- Preeclampsia or gestational hypertension

Effectiveness of the Intervention

Efficacy outcomes of the intervention

Outcome	Metformin (N=110)[a]		Insulin (N=107)		p-value
	Result	n[b]	Result	n[b]	
Pregnancy outcomes					
Gestational age at delivery, gw (mean ± SD)	39.2 ± 1.40	103–109	39.4 ± 1.58	101–107	0.43
Perinatal mortality	NR	NR	NR	NR	NR
Mode of birth	NR	NR	NR	NR	NR
Induction of labour, n (%)	41 (37.6)	103–109	58 (54.2)	101–107	0.014
Assisted vaginal delivery, n (%)	9 (8.3)	103–109	8 (7.5)	101–107	0.83
Instrumental delivery, n (%)	NR	NR	NR	NR	NR
C-section (unspecified if emergency or planned), n (%)	15 (13.8)	103–109	18 (16.8)	101–107	0.53
Maternal gestational weight gain, kg (mean ± SD)	7.97 ± 5.24	103–109	7.82 ± 5.27	101–107	0.83
Preterm delivery	NR	NR	NR	NR	NR
Maternal outcomes					
Maternal wellbeing	NR	NR	NR	NR	NR

Study Reference	Pellonpera 2016, Huhtala 2018					
	Postpartum haemorrhage	NR	NR	NR	NR	NR
	Method of infant feeding					
	Breastfeeding exclusively, months	2.76 ± 2.37	91	2.58 ± 2.43	91	NR[c]
	Breastfeeding altogether, months	6.31 ± 4.00	91	6.59 ± 4.44	91	NR[c]
Post-pregnancy DM	4 (3.9)	102	5 (5.0)	100	NR[c]	
IFG, n (%)						
6–8 weeks postpartum	6 (6.1)	98	4 (4.2)	95	NR[c]	
1 year postpartum	20 (22.5)	89	18 (20.0)	90	NR[c]	
IGT, n (%)						
6–8 weeks postpartum	6 (6.1)	98	1 (1.1)	95	NR[c]	
1 year postpartum	17 (19.1)	89	14 (15.6)	90	NR[c]	
IFG and IGT, n (%)						
6–8 weeks postpartum	1 (1.0)	98	0 (0.0)	95	NR[c]	
1 year postpartum	7 (7.9)	89	6 (6.7)	90	NR[c]	
AGT, n (%)						
6–8 weeks postpartum	12 (12.2)	98	8 (8.4)	95	NR[c]	
1 year postpartum	33 (37.1)	89	28 (31.1)	90	NR[c]	
Gestational hypertension, n (%)	2 (1.8)	103–109	4 (3.7)	101–107	0.44	
Preeclampsia, n (%)	5 (4.6)	103–109	10 (9.3)	101–107	0.17	
Neonatal outcomes						
Macrosomia, n (%)	5 (4.6)	103–109	1 (0.9)	101–107	0.21	
LGA	NR	NR	NR	NR	NR	
Birth weight, g (mean ± SD)	3610 ± 490	103–109	3590 ± 450	101–107	0.78	
Birth weight, SD (mean ± SD)	0.17 ± 1.05	103–109	0.15 ± 0.96	101–107	0.91	
Birth weight <10th percentile, n (%)	12 (11.4)	103–109	9 (8.4)	101–107	0.46	
Birth weight >90th percentile, n (%)	15 (14.3)	103–109	17 (15.9)	101–107	0.74	
Birth injury	NR	NR	NR	NR	NR	
Shoulder dystocia	NR	NR	NR	NR	NR	
Brachial plexus neuropathy	NR	NR	NR	NR	NR	
Neonatal hypoglycaemia	NR	NR	NR	NR	NR	
Admission to NICU	33 (30.1)	103–109	39 (36.4)	101–107	0.36	

Study Reference	**Pellonpera 2016, Huhtala 2018**				
Apgar 5 min, score (mean ± SD)	8.80 ± 1.02	103–109	8.85 ± 0.98	101–107	0.81
IV glucose, n (%)	25 (23.1)	103–109	25 (23.6)	101–107	0.94

[a]Excluding the 21 women in the metformin group who received additional insulin did not change the results (p>0.152 in all comparisons of method of
infant feeding and post-pregnancy DM). [b]Data available for n number of participants. [c]No comparisons between metformin and insulin only.

Study Reference	
Study Reference	**Pellonpera 2016, Huhtala 2018**
	Data on maternal glycaemia (OGTT and HbA1c) at diagnosis of GDM and postpartum, as well as the amino acid profile, were also reported. In addition, the study included data from women achieving fasting glucose <5.5 and postprandial glucose <7.8 mmol/L during diet treatment prior to randomisation.
Authors' Conclusions	Both the women treated by metformin and insulin are at a similarly higher risk of having IGT or diabetes 1 year postpartum. The birth weights between the medical treatment groups were similar. Compared to insulin, metformin treatment of GDM caused a greater increase in amino acids and additional studies and follow-up data are required to ensure the safety of metformin use in GDM pregnancy.

Abbreviations: AGT: abnormal glucose tolerance, including participants with IFG, IGT or DM; DM: diabetes mellitus; GDM: gestational diabetes mellitus; gw: gestational week; HbA1c: glycosylated haemoglobin; IFG: impaired fasting glucose; IGT: impaired glucose tolerance; IV: intravenous; NR: not reported; OGTT: oral glucose tolerance test; SD: standard deviation; WHO: World Health Organisation.

Table 90: Kokic 2018

Study Reference

Characteristic	Experimental Group (n=18)	Control Group (n=20)	p-value
Age, years (mean ± SD)	32.78 ± 3.83	31.95 ± 4.91	0.478
Cardiometabolic health (mean ± SD)			
Body height, m	1.67 ± 0.07	1.68 ± 0.06	0.762

Study Reference	
Kokic 2018	
Study Design	Design A randomised controlled trial Objective To investigate the health-related effects of implementing a supervised, individualised, structured exercise programme, consisting of aerobic and resistance exercises, on the course and outcomes of GDM. Dates Conducted between July 2014 and January 2015 Country Croatia Setting University Hospital Centre Zagreb

Pre-pregnant body mass, kg	68.03 ± 13.65	71.60 ± 15.48	0.515
Pre-pregnant BMI, kg/m^2	24.39 ± 4.89	25.29 ± 4.65	0.515
BMI, kg/m^2	NR	NR	NR

Population recruitment and eligibility	Patient	*Weight, kg*	NR	NR	NR

Characteristics

Recruitment: By direct contact at two university hospitals in Zagreb, Croatia.
Inclusion criteria: An established diagnosis of gestational diabetes, aged between 20 and 40. The upper limit for gestational age at the time of inclusion was set at 30 weeks, to allow a minimum exercise period of 6 weeks, until at least the 36th week of pregnancy.
Exclusion criteria: A medical history of diabetes and miscarriages, pharmacological treatment prior to enrolment in the trial, existing comorbidities, and contraindications for exercise as outlined in criteria published by the ACOG. **Other:** NR

Sample size
N screened/invited = 432
N eligible = 42
N enrolled = 42
N excluded (with reason) = 6 (ineligibility: miscarriage [N=3], twin pregnancy [n=1], other reasons (not provided) [N=2]) N lost to follow-up = 4
N completed = 38
N excluded from analysis = 0 N included in analysis = 38

Maternal demographics **Kokic 2018**

Study Reference

Ethnicity, n (%)			
White	NR	NR	NR
Black	NR	NR	NR
South Asian	NR	NR	NR
East Asian	NR	NR	NR
Mixed	NR	NR	NR
Medical history/risk factors, n (%)			
Hypertension	NR	NR	NR
Diabetes	NR	NR	NR
Pre-pregnant smoking	NR	NR	NR
Pre-pregnant alcohol use	NR	NR	NR
Pre-pregnancy regular physical activity	9 (50.00)	15 (75.00)	0.196
Positive family history of DM	7 (38.89)	8 (40.00)	0.965
Obstetric history, n (%)			
Nulliparous	NR	NR	NR
Parous without GDM	NR	NR	NR
Parous with GDM	NR	NR	NR
Parity (mean ± SD)	0.72 ± 0.83	0.85 ± 0.99	0.806
Education level, n (%)			0.851
Secondary level	7 (38.89)	7 (35.00)	NR
Tertiary level	11 (61.11)	13 (65.00)	NR

Definition of GDM

Criteria published by the International Association of the Diabetes and Pregnancy Study Groups

Study Reference **Kokic 2018**

Methods

Duration of follow-up

Until assessment immediately following childbirth

Method of assigning treatment arm

Participants were randomised by block randomisation using a web-based computerized procedure into two groups: experimental and control. The staff involved with the exercise sessions and assessments had no influence on the randomisation procedure. Due to the nature of the study, participants were not blinded. Physicians and laboratory staff were blinded.

Arm 1 (n=18)

Experimental Group

Women were started on an individualised, structured exercise programme two times per week, along with their standard prenatal care. Participants in this group were also asked to undertake at least 30 min of brisk walking per day. In addition, all participants were commenced on medical nutrition therapy recommended for women with GDM. Full details of the exercise regime and the nutritional therapy are reported in the publication. Arm 2 (n=20)

Control Group

Women received standard prenatal care for GDM alone, but were not discouraged from exercising on their own. In addition, all participants were commenced on medical nutrition therapy recommended for women with GDM, full details of which are reported in the publication.

Outcomes

Primary endpoint

- Fasting and postprandial glucose levels as measured monthly or bimonthly for the duration of pregnancy
- Neonatal anthropometric data, including body mass, arm circumference and skinfold thickness

Secondary endpoints

- Body mass and gain during pregnancy
- Activity levels as measured by the Pregnancy Physical Activity Questionnaire

Outcome	Experimental Group (n=18)	Control Group (n=20)	p-value
Pregnancy outcomes			
Gestational age at birth, gw (mean ± SD)	38.89 ± 0.90	39.45 ± 0.60	0.063
Perinatal mortality	NR	NR	NR
Mode of birth			
Induction of labour, n (%)	(11.11)	(35)	0.346
Vaginal delivery, n (%)	NR	NR	NR
Instrumental delivery, n (%)	(5.56)	(0)	0.784

Study Reference

	C-section (unspecified if emergency or planned), n (%)	(27.78)	(25)	0.696
Effectiveness of Efficacy outcomes of the intervention **the Intervention Kokic 2018**	*Prolonged labour, n (%)*	(5.56)	(10)	0.633
	Maternal gestational weight gain	NR	NR	Not significant
	Preterm delivery	All subjects gave birth between the 38th and 40th week of pregnancy		
	Maternal outcomes			
	Maternal wellbeing	NR	NR	NR
	Postpartum haemorrhage	NR	NR	NR
	Method of infant feeding	NR	NR	NR
	Post-pregnancy T2DM	NR	NR	NR
	Pregnancy-induced hypertension	NR	Two cases, one of which progressed to preeclampsia	NR
	Neonatal outcomes			
	Macrosomia	NR	NR	NR
	LGA	NR	NR	NR
	Birth injury	NR	NR	NR
	Shoulder dystocia	NR	NR	NR
	Brachial plexus neuropathy	NR	NR	NR
	Neonatal hypoglycaemia, n (%)	0 (0)	0 (0)	1.000
	Admission to NICU	NR	NR	NR
	Other neonatal complications (hyperbilirubinaemia), n (%)	0 (0)	1 (5)	0.806
	Apgar 1 min, score (mean ± SD)			
	Apgar 5 min, score (mean ± SD)			
	Neonatal body mass, g (mean ± SD)	3514.45 ± 413.57	3377.00 ± 494.27	0.393
	Neonatal length, cm (mean ± SD)	50.11 ± 2.25	50.25 ± 2.51	0.851
	Neonatal PI, kg/m^3 (mean ± SD)	2.66 ± 0.63	2.65 ± 0.16	0.093
	Neonatal BMI, kg/m^2 (mean ± SD)	13.96 ± 0.97	13.21 ± 1.01	0.035

Physical activity and glucose levels (fasting and postprandial) at the end of pregnancy were also reported in detail in the publication.

Authors' Conclusions	The structured exercise programme had a beneficial effect on postprandial glucose levels at the end of pregnancy. However, the exercise programme did not reduce the rate of complications during birth. Neonatal body mass index was slightly higher in the experimental group (p=0.035), but still well within healthy limits.

Abbreviations: ACOG: American College of Obstetricians and Gynaecologists; DM: diabetes mellitus; gw: gestational week; NR: not reported; SD: standard deviation.

Table 91: MFMU Network RCT, Palatnik 2015, Casey 2015

Study Reference	**MFMU Network RCT, Palatnik 2015, Casey 2015**
Study Design	Design Randomised controlled trial Objective To determine whether there is an association between gestational age at the time of treatment initiation for mild GDM and perinatal outcomes. To determine whether maternal BMI might alter the impact of therapy on foetal growth in women with mild GDM. Dates Between October 2002 and November 2007. Country US Setting The Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units (MFMU) Network

Population Patient recruitment and eligibility **Characteristics**

Recruitment: NR

Inclusion criteria: Pregnant women screened between 24 weeks 0 days and 30 weeks 6 days gestation and diagnosed with mild GDM.

Exclusion criteria: OGTT fasting values above 95 mg/dL. **Other:** NR

Sample size
N screened/invited = 19,665
N eligible = 958
N enrolled = 958

Study Reference

N excluded (with reason) = NR
N lost to follow-up = NR
N completed = NR
N excluded from analysis = 27 (due to missing delivery data)

Characteristic	Treatment (n=485)	Usual care (n=473)	p-value
Age, years (mean ± SD)	29 ± 5.1	29 ± 5.6	0.86
Cardiometabolic health			
Pre-pregnant BMI, kg/m^2	NR	NR	NR
BMI category, kg/m^2			1.0

N included in analysis = 932

Maternal demographics

Study Reference	MFMU Network RCT, Palatnik 2015, Casey 2015		
Normal weight (< 25), n (%)	73 (15)	70 (15)	NR
Overweight (25–29.9), n (%)	187 (38)	181 (38)	NR
Class I Obese (30–34.9), n (%)	153 (32)	151 (32)	NR
Class II Obese (35–39.9), n (%)	53 (11)	57 (11)	NR
Class III Obese (≥ 40), n (%)	19 (4)	20 (4)	NR
50g GCT, mg% (mean ± SD)	159 ± 15.3	160 ± 15.5	0.50
OGTT fasting, mg% (mean ± SD)	87 ± 5.5	86 ± 5.7	0.34
OGTT 1h, mg% (mean ± SD)	192 ± 21.9	193 ± 19.3	0.11
OGTT 2h, mg% (mean ± SD)	174 ± 21.8	173 ± 19.6	0.84
OGTT 3h, mg% (mean ± SD)	137 ± 29.0	134 ± 31.5	0.14
Weight, kg	NR	NR	NR
Ethnicity, n (%)			0.55
White	123 (25)	119 (25)	NR
Black	56 (12)	54 (12)	NR
Hispanic	281 (58)	265 (56)	NR
South Asian	NR	NR	NR
East Asian	NR	NR	NR
Other	25 (5)	35 (7)	NR
Medical history/risk factors, n (%)			
Hypertension	NR	NR	NR
Diabetes	NR	NR	NR
Pre-pregnant smoking	Only reported by gestational week at randomisation.		0.97
Pre-pregnant alcohol use	Only reported by gestational week at randomisation.		0.14
Gestational age at randomisation to treatment, weeks (mean ± SD)	29 ± 0.6	29 ± 1.5	0.13
Obstetric history, n (%)			
Primigravida	104 (21)	123 (26)	0.10
Parous without GDM	NR	NR	NR
Parous with GDM	NR	NR	NR
Education level	NR	NR	NR

Definition of GDM
50g GCT with a 1-hour blood glucose value between 135–200 mg/dL followed by 3-hour OGTT. Mild GDM defined as a fasting blood glucose <95 mg/dL and ≥2 post-challenge glucose above the following thresholds: 1-hour >180 mg/dL, 2-hour >155 mg/dL, 3-hour >140 mg/dL.

Study Reference	MFMU Network RCT, Palatnik 2015, Casey 2015
Methods	Duration of follow-up Until shortly after birth. Method of assigning treatment arm Randomisation at each clinical centre. Arm 1 (n=485) **Treatment** Treatment included formal nutritional counselling and diet therapy. Insulin was prescribed if the majority of fasting or postprandial values met or exceeded 95 mg% and 120 mg% respectively. Details of how the intervention was delivered NR. Arm 2 (n=473) **Usual prenatal care** Details of how the intervention was delivered NR. Outcomes **Primary endpoint** • A composite outcome that included perinatal mortality and complications that have been associated with maternal hyperglycaemia: neonatal hypoglycaemia, defined as a glucose value of less than 35mg/dl; hyperbilirubinemia, defined as bilirubin value greater than the 95th percentile for any given point after birth; hyperinsulinemia, defined as a cord-blood C-peptide level greater than the 95th percentile and birth trauma, defined as brachial plexus palsy or clavicular, humeral, or skull fracture. **Secondary endpoints** • Occurrence of LGA, defined as birth weight above the 90th percentile of a US reference population. • NICU admission. • Gestational hypertension/preeclampsia. All cases of hypertensive disorders underwent masked central review by two of the investigators to ensure accurate diagnosis. • Caesarean delivery. • Foetal hyperinsulinemia, defined as umbilical cord c-peptide level >1.77 ng/mL corresponding to 95th percentile from an unselected obstetrical population of women in the MFMU Network of the Eunice Kennedy Shriver National Institute of Child Health and Human Development. • Neonatal fat mass, calculated according to a published method based on the infant's length, head and upper mid-arm circumferences, and flank skinfold.

Study Reference **MFMU Network RCT, Palatnik 2015, Casey 2015**

Efficacy outcomes of the intervention

Outcome	24–26 weeks		27–29 weeks[a]		p-value
	Treatment (n=69)[b]	Usual care (n=43)[b]	Treatment (n=288)[b]	Usual care (n=282)[b]	
Pregnancy outcomes					
Gestational age at birth	NR	NR	NR	NR	NR
Perinatal mortality	NR	NR	NR	NR	NR
Mode of birth	NR	NR	NR	NR	NR
Induction of labour	NR	NR	NR	NR	NR
Vaginal delivery	NR	NR	NR	NR	NR

[a]Women who received intervention at ≥30 week of gestation were excluded due to much smaller window for therapeutic intervention. [b]Some of the denominators in each outcome are smaller than the n due to missing delivery data. [c]Shoulder dystocia was not included in the analysis as there were only 25 cases.
[d]Included perinatal mortality and complications that have been associated with maternal hyperglycaemia: neonatal hypoglycaemia, hyperbilirubinemia, hyperinsulinemia, and birth trauma.

Study Reference	MFMU Network RCT, Palatnik 2015, Casey 2015				
Instrumental delivery	NR	NR	NR	NR	NR
C-section (unspecified if emergency or planned), n (%)	23 (33.8)	15 (34.9)	77 (26.7)	93 (33.0)	0.57
aternal gestational weight in	NR	NR	NR	NR	NR
eterm delivery	NR	NR	NR	NR	NR
aternal outcomes	NR	NR	NR	NR	NR
aternal wellbeing	NR	NR	NR	NR	NR
stpartum haemorrhage	NR	NR	NR	NR	NR
ethod of infant feeding	NR	NR	NR	NR	NR
st-pregnancy T2DM	NR	NR	NR	NR	NR
stational pertension/preeclampsia, n)	7 (10.3)	6 (14.0)	26 (9.0)	37 (13.1)	0.91
onatal outcomes	NR	NR	NR	NR	NR
acrosomia	NR	NR	NR	NR	NR
A, n (%)	8 (11.6)	6 (14.0)	20 (6.9)	40 (14.2)	0.36
th injury	NR	NR	NR	NR	NR
oulder dystocia[c]	NR	NR	NR	NR	NR
achial plexus neuropathy	NR	NR	NR	NR	NR
onatal hypoglycaemia	NR	NR	NR	NR	NR
mission to NICU, n (%)	10 (14.5)	7 (16.3)	25 (8.7)	38 (13.5)	0.55
mposite outcome, n (%)[d]	25 (37.3)	20 (50.0)	90 (32.3)	93 (33.9)	0.32

- There was no significant interaction between gestational age at randomisation and treatment group for birth weight Z-scores based on a U.S. reference population (p-value of 0.86 for the interaction of gestational age and treatment group)
- There were significant reductions in LGA birth weight in treated women with a BMI between 25 and 40 kg/m^2 at enrolment, as incidence of LGA among Class I obese women who received treatment was 13 (9%), compared to 29 (20%) for women receiving usual prenatal care (p=0.005). No such effect of treatment was evident in women who were of normal weight or morbidly obese.
- Women in the lowest (< 25 kg/m^2) or highest (≥ 40 kg/m^2) BMI categories delivered infants with similar neonatal fat mass regardless of treatment assignment. In contrast, neonatal fat mass was significantly reduced with diet therapy and routine glucose monitoring in women with a BMI between 25–40 kg/m^2. Overweight treated women delivered infants with a mean fat mass of 404 ± 189 g compared to 455 ± 210 g for women who received routine care (p-value NR).
- Umbilical cord serum c-peptide was elevated in 20% of the entire cohort. However, consideration of maternal BMI at enrolment did not modify the small but statistically insignificant treatment effect previously reported from this randomized trial (p=0.16).

Study Reference	MFMU Network RCT, Palatnik 2015, Casey 2015

Study Reference	MFMU Network RCT, Palatnik 2015, Casey 2015

- Additional data presented in the publications include: results by gestational age at randomisation for each individual week, plots of odds ratios for treatment versus control group by gestational age at randomisation for the outcomes of interest and plots of outcomes according to maternal BMI category and treatment group.

Authors' Conclusions	Earlier initiation of treatment of mild GDM within the recommended gestational age range for screening was not associated with stronger effect of treatment on perinatal outcomes. There was a beneficial effect of treatment on foetal growth in women with mild GDM who were overweight or Class I and II obese. These effects were not apparent for normal weight and very obese women.

Abbreviations: BMI: body mass index; GDM: gestational diabetes mellitus; LGA: large for gestational age; NICU: neonatal intensive care unit; NR: not reported.

SLRs

Table 92: Brown 2017L, Lifestyle SLR

Study Reference	Brown 2017L (Lifestyle SLR)
Study Design	Design Systematic literature review Objective To evaluate the effects of combined lifestyle interventions with or without pharmacotherapy in treating women with gestational diabetes. Search dates NR Country Various Setting NR

Study eligibility

Population	Pregnant women diagnosed with gestational diabetes (diagnosis as defined by the individual trial). Women with known type 1 or type 2 diabetes were excluded

Study Reference **Brown 2017L (Lifestyle SLR)**

Inclusion (PICOS)

Population Characteristics	
Intervention	Lifestyle interventions which could include a combination of at least two or more of the following: diet, physical activity, education, behavioural change techniques, regimens of self-monitoring of blood glucose, or any other intervention, with or without pharmacotherapy.
Comparator	Standard care or another lifestyle intervention
Outcomes	**Primary** *Maternal* • Hypertensive disorders of pregnancy (including preeclampsia, pregnancy-induced hypertension, eclampsia as defined by trialists • Caesarean section • Development of type 2 diabetes *Neonatal* • Perinatal (fetal and neonatal death) and later infant mortality • Large-for-gestational age (LGA) (as defined by trialists) • Death or serious morbidity composite (variously defined by trials, e.g. perinatal or infant death, shoulder dystocia, bone fracture or nerve palsy) • Neurosensory disability in later childhood (as defined by trialists) **Secondary** *Maternal* Use of additional pharmacotherapy; maternal hypoglycaemia (as defined by trialists); glycaemic control during/end of treatment (as defined by trialists); weight gain in pregnancy; adherence to the intervention; induction of labour; placental abruption; postpartum haemorrhage (as defined by trialists); postpartum infection; perineal trauma/tear; breastfeeding at discharge, six weeks postpartum, six months or longer; maternal mortality; sense of well-being and quality of life; behavioural changes associated with the intervention; views of the intervention; relevant biomarker changes associated with the intervention (including adiponectin, free fatty acids, triglycerides, high-density lipoproteins (HDL), low-density lipoproteins (LDL), insulin) *Long-term outcomes for mother* Postnatal depression; body mass index (BMI); postnatal weight retention or return to pre-pregnancy weight; type 1 diabetes; impaired glucose tolerance; subsequent gestational diabetes; cardiovascular health (as defined by trialists including blood pressure, hypertension, cardiovascular disease, metabolic syndrome)

Study Reference **Brown 2017L (Lifestyle SLR)**

	Fetal/neonatal outcomes Stillbirth; neonatal death; macrosomia (greater than 4000 g; or as defined by individual study); small-for-gestational (SGA) age (as defined by trialists); birth trauma (shoulder dystocia, bone fracture, nerve palsy); gestational age at birth; preterm birth (< 37 weeks' gestation; and < 32 weeks' gestation); five-minute Apgar less than seven; birthweight and z score; head circumference and z score; length and z score; ponderal index; adiposity (including skinfold thickness measurements (mm); fat mass as defined by trialists); neonatal hypoglycaemia (as defined by trialists); respiratory distress syndrome; neonatal jaundice (hyperbilirubinaemia) (as defined by trialists); hypocalcaemia (as defined by trialists); polycythaemia (as defined by trialists); relevant biomarker changes associated with the intervention (including insulin, cord c-peptide) *Later infant/childhood outcomes* Weight and z score; height and z score; head circumference and z score; adiposity (including BMI, skinfold thickness, fat mass); educational attainment; blood pressure; type 1 diabetes; type 2 diabetes; impaired glucose tolerance; dyslipidaemia or metabolic syndrome. *Child as an adult outcomes* Weight; height; adiposity (including BMI, skinfold thickness, fat mass); employment, education and social status/achievement; dyslipidaemia or metabolic syndrome; type 1 diabetes; type 2 diabetes; impaired glucose tolerance; cardiovascular health (as defined by trialists including blood pressure, hypertension, cardiovascular disease, metabolic syndrome) *Health service use* Number of antenatal visits or admissions; number of hospital or health professional visits (including midwife, obstetrician, physician, dietician, diabetic nurse) ; admission to neonatal intensive care unit/nursery; duration of stay in neonatal intensive care unit or special care baby unit; length of antenatal stay; length of postnatal stay (maternal); length of postnatal stay (baby); cost of maternal care; cost of offspring care; costs associated with the intervention; costs to families associated with the management provided; cost of dietary monitoring (e.g. diet journals, dietician, nurse visits, etc); costs to families - change of diet, extra antenatal visits; extra use of healthcare services (consultations, blood glucose monitoring, length and number of antenatal visits); women's view of treatment advice
Study design	Published or unpublished randomised controlled trials in full-text or abstract format

Exclusion (reasons given in excluded study list)

- Women with type 1 or type 2 diabetes
- Non-RCTs or quasi-randomised trials
- Cross-over trials
- Interventions examining the comparison of different dietary interventions or the effects of exercise alone (included in other Cochrane systematic reviews)

Other NR Flow of Studies (PRISMA)

- Database results: 253
- Hand-searches/other sources: 21
- 274 records after duplicates removed

Study Reference **Brown 2017L (Lifestyle SLR)**

- Title/abstracts reviewed: 73
- Full-texts reviewed: 73
- Articles included in qualitative synthesis: 25
- Articles included in quantitative synthesis (meta-analysis): 15 studies associated with 45 publications

Included study characteristics

Characteristic	Details
Design	All 15 included studies were RCTs
Sample sizes	Minimum 19 – maximum 1000 participants 12 studies had sample size of ≤300
Setting and timing	USA (N=4; China (N=2); Iran (N=2); Canada (N=2); UK (N=1); Italy (N=1): UAE (N=1); Thailand (N=1); Australia and UK (N=1)
Participants	Maternal age (reported by 11 trials): mean age in intervention group – minimum of 29.2 (SD 5.7) years to maximum 35.9 (SD 4.8) years; mean age in control group – minimum of 28.9 (SD 5.6) to 33.9 (SD 5.3) years BMI (reported by 7 trials): mean BMI in intervention group – minimum of 22.9 (SD 3.6) to maximum 31.2 (SD 6.7); mean BMI in control group – minimum of 23.4 (SD 3.9) to maximum of 30.2 (SD 5.1) Ethnicity (reported by 9 trials)
Diagnostic criteria for GDM	Six different diagnostic criteria were used in the 9 trials where details were provided: • World Health Organization (1999): 3 trials • Carpenter and Coustan criteria: 2 trials • American Diabetes Association (2000): 1 trial • ADIPS (Hofman 1998): 1 trial • IADPSG criteria: 1 trial • Hatem (1988) 75g OGTT >7.5 mmol (second trimester) and >9.6 mmol/L (third trimester), no other details: 1 trial • 6 trials did not provide details on the criteria used to diagnose the women with gestational diabetes
Treatment targets	NR

Study Reference	Brown 2017L (Lifestyle SLR)

	Interventions and comparisons	**Bancroft 2000**: Intensive intervention (standard dietary advice, glucose monitoring five days a week, HbA1c monthly, serial ultrasound, Doppler studies, cardiotocography (CTG monitoring) compared with usual care (dietary advice, HbA1c monthly). **Bo 2014**: Reported on a multiple-arm trial that included a) Individualised- dietary advice alone, b) Exercise alone, c) Behavioural intervention and d) Behavioural intervention and exercise. We used the combined behavioural and exercise group as the intervention arm for this review and the Individualised-dietary advice alone as the control group.	
		Crowther 2005: Intensive intervention (individualised-dietary advice, advice on self-monitoring of blood glucose) compared with usual care (women and caregivers unaware of diagnosis). **Elnour 2008**: Intensive intervention (structured pharmaceutical care, structured education, self-monitoring of blood glucose) compared with usual care (no additional education or pharmacist counselling). **Ferrara 2011**: Intensive intervention (individualised advice on diet, exercise and breastfeeding) compared with usual care (printed material only in prenatal and postnatal period). **Garner 1997**: Intensive intervention (dietary counselling, self-glucose monitoring, biweekly review, monitoring of fetal growth, amniotic volume and cardiac size) versus usual care (no dietary counselling). **Gillen 2004**: Group session on education and diet followed by specific dietary advice compared with group session on education and diet followed by standard clinical care and advice. **Jovanovic-Peterson 1989**: Diet alone compared with diet plus supervised exercise. **Kaviani 2014**: Relaxation training (education, breathing, muscle relaxation, mental imagery, and contacted by telephone by the researcher three times per week) compared with usual care (no details). **Landon 2009**: Nutritional counselling and diet therapy +/- insulin plus self-monitoring of blood glucose compared with usual care +/- insulin plus self-monitoring of blood glucose. **Mendelson 2008**: Intensive education and spiritual intervention compared with standard education. **Rahimikian 2014**: Face-to-face education (risks of GDM, training on glycaemic control, exercise, diet, medication and follow-up) compared with usual care (no details). **Yang 2003**: Intensive intervention (including diet and exercise advice, self-monitoring of blood glucose, insulin if required, fortnightly specialist review) versus usual care (no details). **Yang 2014**: Shared care protocol adapted from Crowther 2005. Individualised and group dietary and physical activity counselling, self-monitoring blood glucose compared with usual care (group education on exercise and physical activity, not specifically taught blood glucose self-monitoring). **Youngwanichsetha 2014**: Mindfulness eating and yoga compared with standard diabetes care (no details).	

Study Reference **Brown 2017L (Lifestyle SLR)**

Outcomes	**Maternal primary outcomes** • Pregnancy-induced hypertension: 4 trials • Caesarean section: 10 trials • Development of type 2 diabetes: two trials **Neonatal primary outcomes** • LGA: 6 trials • Perinatal death: 2 trials • Composite of serious neonatal outcomes: 3 trials Data were also available for the following maternal **secondary outcomes**: need for supplementary medication, maternal hypoglycaemia, fasting plasma glucose concentration, postprandial glucose concentration, HbA1c, weight gain in pregnancy, induction of labour, postpartum haemorrhage, postnatal infection/pyrexia, perineal trauma/tear, breastfeeding, postnatal depression, quality of life, impaired glucose tolerance, metabolic syndrome and return to pre-pregnancy weight. Data were available for the following neonatal secondary outcomes: stillbirth, neonatal death, macrosomia, small-for-gestational age (SGA), birth trauma (shoulder dystocia, bone fracture, nerve palsy), gestational age at birth, preterm birth, congenital anomaly, five-minute Apgar less than seven, birthweight, length, neonatal fat

Study Reference	Brown 2017L (Lifestyle SLR)
	mass, neonatal hypoglycaemia, respiratory distress syndrome, hyperbilirubinaemia, hypocalcaemia, polycythaemia, childhood growth, childhood cholesterol and childhood impaired glucose tolerance. Data were available for the following health service outcomes: visits to health professionals, antenatal hospital admissions and admission to neonatal intensive care unit.
Funding	Reported in 7 trials. None of the sources were conditional grants from pharmaceutical companies. The remaining trials did not detail the sources of funding (if any) in the published manuscript
Conflicts of interest	Declarations of interest were made in 4 trials. • 3 reported that there were no conflicts of interest for any of the authors • 1 trial reported that there was a conflict of interest for one of the 12 authors. The conflict states that the authors institution had received research funding from Eli Lilly and the author is a member of advisory committee and speaker forum sponsored by Eli Lilly. • The remaining trials did not provide any statements about conflict of interest.

Definition of GDM

Diagnosis as defined by the individual trial

Study Reference **Brown 2017L (Lifestyle SLR)**

Searches Sources searched:

Cochrane Pregnancy and Childbirth's Trials Register, which is maintained by an Information Specialist and contains trials identified from:

- CENTRAL (monthly searches)
- MEDLINE (weekly searches via Ovid)
- Embase (weekly searches via Ovid)
- CINAHL (monthly searches via EBSCO)
- Hand-searches of 30 journals and the proceedings of major conferences
- Weekly current awareness alerts for a further 44 journals plus monthly BioMed Central email alerts

Screening and selection process

Search results were screened by two people and the full text of all relevant trial reports identified through the searches were reviewed. Based on the intervention described, each trial report was assigned a number that corresponds to a specific Pregnancy and Childbirth review topic (or topics), and was then added to the register. Two review authors independently assessed for inclusion all the potential studies identified as a result of the search strategy. Any disagreement was resolved through discussion, or if required, through consultation with a third person.

Study quality assessment

Two reviewers independently assessed risk of bias for each randomised study using the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions. Any disagreement was resolved by discussion or by involving a third assessor.

Methods for combining intervention evidence

A form was designed to extract data. For eligible studies, two review authors extracted the data using the agreed form. Discrepancies were resolved through discussion or, if required, a third person was consulted. Data were entered into Review Manager software (RevMan 2014) and checked for accuracy. When information regarding any of the above was unclear, contact authors of the original reports were contacted in attempt to provide further details. The quality of the evidence was assessed using the GRADE approach. Up to a maximum of seven outcomes were selected for the mother and seven for the infant, covering both short- and long-term outcomes for the main comparisons. For dichotomous data, results were presented as summary risk ratio with 95% confidence intervals. For continuous data, the mean difference was used if outcomes were measured in the same way between trials. The standardised mean difference was planned to be used to combine trials that measured the same outcome, but using different methods. Statistical analysis was carried out using the Review Manager software (RevMan 2014). A fixed-effect meta-analysis was used for combining data where it was reasonable to assume that studies were estimating the same underlying treatment effect: i.e. where trials examined the same intervention, and the trials' populations and methods were judged sufficiently similar. If clinical heterogeneity was sufficient to expect that the underlying treatment effects differed between trials, or if substantial statistical heterogeneity was detected, a random-effects meta-analysis was produced for an overall summary, if an average treatment effect across trials was considered clinically meaningful. The random-effects summary was treated as the average of the range of possible treatment effects and the clinical implications of treatment effects differing between trials were discussed. If the average treatment effect was not clinically meaningful, trials were not combined. Where random-effects analyses were used, the results were presented as the average treatment effect with 95%confidence intervals, and the estimates of Tau^2 and I^2.

Study Reference **Brown 2017L (Lifestyle SLR)**

Efficacy outcomes for lifestyle interventions versus usual care or diet alone for the treatment of women with gestational diabetes

	Outcome	Anticipated absolute effects[a] (95% CI) Risk with usual care/control	Anticipated absolute effects[a] (95% CI) Risk with lifestyle intervention	Risk ratio (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
Effectiveness of the Intervention	**Pregnancy outcomes**					
	Gestational age at birth	NR	NR	NR	NR	NR
	Hypertensive disorders of pregnancy (pre-eclampsia)	129 per 1000	90 per 1000 (51-157)	0.70 (0.40-1.22)	2796 (4 RCTs)	Low
	Perinatal mortality (fetal and neonatal death) and later infant mortality	5 per 1000	0 per 1000 (0-9)	0.09 (0.01-1.70)	1988 (2 RCTs)	Low
	Mode of birth					
	Induction of labour	211 per 1000	252 per 1000 (220-285)	1.20 (0.99-1.46	2699 (4 RCTs)	High
	Vaginal delivery	NR	NR	NR	NR	NR
	Instrumental delivery	NR	NR	NR	NR	NR
	C-section (not specified if planned or emergency)	380 per 1000	342 per 1000 (296-399)	0.90 (0.78-1.05)	3545 (10 RCTs)	Low
	Maternal gestational weight gain	NR	NR	NR	NR	NR
	Preterm delivery	NR	NR	NR	NR	NR
	Perineal trauma/tear	498 per 1000	518 per 1000 (463-588)	1.04 (0.93-1.18)	1000 (1 RCT)	Moderate
	Maternal outcomes					
	Maternal wellbeing	NR	NR	NR	NR	NR
	Postpartum haemorrhage	NR	NR	NR	NR	NR
	Method of infant feeding	NR	NR	NR	NR	NR
	Postnatal depression	169 per 1000	83 per 1000 (53 - 132)	0.49 (0.31–0.78)	573 (1 RCT)	Low

Study Reference	Brown 2017L (Lifestyle SLR)					
	Postnatal weight retention or return to pre-pregnancy weight	214 per 1000	375 per 1000 (225-621)	1.75 (1.05-2.90)	156 (1 RCT)	Low
	Post-pregnancy T2DM	83 per 1000	81 per 1000 (45-146)	0.98 (0.54-1.76)	486 (2 RCTs)	Low
	Neonatal outcomes					
	Macrosomia	NR	NR	NR	NR	NR
	LGA	189 per 1000	113 per 1000 (95-134)	0.60 (0.50–0.71)	2994 (6 RCTs)	Moderate
	Composite outcome (death, shoulder dystocia, nerve palsy, bone fracture)	193 per 1000	110 per 1000 (41 to 299)	0.57 (0.21-1.55)	1930 (2 RCTs)	Very low
	Adiposity (neonatal) – neonatal fat mass (g)	The mean neonatal fat mass was 427 g	Mean neonatal fat mass in the intervention group was 37.30 g fewer (63.97 fewer to 10.63 fewer)	–	958 (1 RCT)	Low
	Shoulder dystocia	NR	NR	NR	NR	NR
	Brachial plexus neuropathy	NR	NR	NR	NR	NR
	Neonatal hypoglycaemia	75 per 1000	74 per 1000 (49-114)	0.99 (0.65-1.52)	3000 (6 RCTs)	Moderate
	Admission to NICU	NR	NR	NR	NR	NR

t	Random sequence generator (selection bias)	Allocation concealment (selection bias)	) Blinding of participants and personnel performance bias	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	) Selective reporting reporting bias	Other bias
Bancroft 2000	❑	❑	❑	❑	❑	❑	❑
Bo 2014							

Study Reference	Brown 2017L (Lifestyle SLR)						
Crowther 2005							
Elnour 2008							
Ferrara 2011							
Garner 1997							
Gillen 2004							
Jovanovic-Peterson 1989							
Kaviani 2014							
Landon 2009							
Mendelson 2008							
Rahimikian 2014							
Yang 2003							
Yang 2014							
Youngwanichsetha 2014							

Women receiving lifestyle interventions were less likely to have postnatal depression and were more likely to achieve postpartum weight goals. Exposure to lifestyle interventions was associated with a decreased risk of the baby being born LGA and decreased neonatal adiposity. Long-term maternal and childhood/ adulthood outcomes were poorly reported. The contribution of individual components of lifestyle interventions could not be assessed. Ten per cent of participants also received some form of pharmacological therapy. Lifestyle interventions are useful as the primary therapeutic

strategy and most commonly include healthy eating, physical activity and self-monitoring of blood glucose concentrations. Future research could focus

s on which specific interventions are most useful (as the sole intervention without pharmacological treatment), which health professionals should give them and the optimal format for providing the information. Evaluation of long-term outcomes for the mother and her child should be a priority when planning future trials. There has been no in-depth exploration of the costs 'saved' from reduction in risk of LGA/macrosomia and potential longer-term risks for the infants.

Study Reference	**Brown 2017L (Lifestyle SLR)**
Design Systematic literature review Objective To evaluate the effects of insulin in treating women with GDM Search dates 1 May 2017 Country Setting	

Table 93: Brown 2017I, Insulin SLR

Study Reference	**Brown 2017I** (Insulin SLR)

Study Reference **Brown 2017I** (Insulin SLR)

Study eligibility **Selection criteria (PICOS):**

RCTs (including those published in abstract form) comparing:

a) Insulin with an oral antidiabetic pharmacological therapy;
b) With a non-pharmacological intervention;
c) Different insulin analogues;
d) Different insulin regimens for treating women diagnosed with GDM

Population	Pregnant women diagnosed with GDM (diagnosis as defined by the individual trial)
Intervention and Comparator	• Insulin (any type) vs oral antidiabetic agents (main comparison) • Insulin type A vs insulin type B (e.g. rapid-acting vs short-acting; intermediate-acting vs long-acting) • Insulin (any type) vs diet/standard care • Insulin (any type) vs exercise • Insulin (any type) vs diet plus exercise • Insulin regimen A vs insulin regimen B • Insulin (any type) vs other treatment intervention not previously described
Outcomes	**Primary** *Maternal* • Hypertensive disorders of pregnancy (including pre-eclampsia, pregnancy-induced hypertension, eclampsia, as defined by trialists) • Caesarean section • Development of type 2 diabetes (as defined by trialists, including results of postnatal testing) *Neonatal* • Perinatal (foetal and neonatal death) and later infant mortality • LGA (as defined by trialists) • Death or serious morbidity composite (as defined by trialists e.g. perinatal or infant death, shoulder dystocia, bone fracture or nerve palsy) • Neurosensory disability in later childhood (as defined by trialists) **Secondary** • Maternal (use of additional pharmacotherapy; maternal hypoglycaemia; glycaemic control during/end of treatment; weigh gain in pregnancy; adherence to the intervention; induction of labour; placental abruption; postpartum haemorrhage; postpartum infection; perineal trauma/tearing; breastfeeding at discharge, six weeks postpartum, six months or longer; maternal mortality; sense of wellbeing and quality of life; behavioural changes associated with the intervention; views of the intervention; relevant biomarker changes associated with the intervention [adiponectin, free fatty acids, triglycerides, high-density lipoproteins, low-density lipoproteins]) • Long-term maternal (postnatal depression; BMI; postnatal weight retention or return to pre-pregnancy weight; type 1 diabetes; type 2 diabetes; impaired glucose tolerance; CV health [blood pressure, hypertension, CVD, metabolic syndrome]) • Foetal/neonatal (stillbirth; neonatal death; macrosomia; SGA; birth trauma [shoulder dystocia, bone fracture, nerve palsy]); gestational age at birth; preterm birth [<37 and <32 weeks gestation]; five-minute Apgar <7; birthweight

Study Reference	**Brown 2017I** (Insulin SLR)
	and z score; head circumference and z score; length and z score; ponderal index; adiposity [including skinfold thickness measurements, fat mass]; neonatal hypoglycaemia; respiratory distress syndrome; neonatal jaundice [hyperbilirubinemia]; hypocalcaemia; polycythaemia; relevant biomarker changes associated with the intervention [insulin, cord c-peptide]) • Later infant/childhood (weight and z scores; height and z scores; head circumference and z scores; adiposity [BMI, skinfold thickness, fat mass], educational attainment; blood pressure; type 1 diabetes; type 2 diabetes; impaired glucose tolerance; dyslipidaemia or metabolic syndrome) • Child as an adult (weight, height, adiposity [BMI, skinfold thickness, fat mass; CV health; employment, education and social status/achievement; dyslipidaemia or metabolic syndrome; type 1 diabetes; type 2 diabetes; impaired glucose tolerance) • Health service use (number of antenatal visits or admissions; number of hospital or health professional visits; admission to NICU/nursery; duration of stay in NICU or special care baby unit; length of antenatal stay; length of postnatal stay (maternal and baby); cost or maternal or offspring care; costs associated with the intervention; costs to families associated with the management provided; cost of dietary monitoring; extra use of healthcare services; women's view of treatment advice)
Study design	RCTs

Study Reference **Brown 2017I** (Insulin SLR)

Exclusion criteria (reasons given in excluded study list):

Quasi-randomised and trials including women with pre-existing type 1 or type 2 diabetes; cross-over trials.

Other:

NR

Flow of studies (PRISMA)

Characteristic	Details
Design	All 53 included studies were RCTs
Sample sizes	Minimum 10 – maximum 733 participants 34 studies had sample size of ≤100
Setting	USA (N=16); India (N=7); Iran (N=6); Egypt (N=3); Brazil (N=3); Pakistan (N=3); Finland (N=3); Italy (N=2); Sweden (N=1); Canada (N=1); Ghana (N=1); Australia (N=1); New Zealand and Australia (N=1); Turkey (N=1); Israel (N=1); Malaysia (N=1); South Africa (N=1); Poland (N=1)
Timing	2010s (N=13); 2000s (N=10); 1990s (N=2); 1990/80s (N=1); 1980s (N=1); 1970s (N=1); 1960s (N=1); 1950s (N=1); no details (N=23)

Study Reference	**Brown 2017I** (Insulin SLR)		
	Participants	Maternal age (years)	**Insulin vs oral antidiabetic** *Insulin vs metformin (N=19)* Insulin arm range: 23.4 ± 2.5 (n=50) to 35 (30–38) (n=43)

Titles/abstracts reviewed = 288
Full texts reviewed = 153
Articles included in qualitative synthesis = 53 studies associated with 103 publications
Articles included in quantitative synthesis (meta-analysis) = 51 studies associated with 99 publications
Included study characteristics - summary

- The 53 included studies reported data for 7381 women and 46 of these studies reported data for 6435 infants (7 studies reported no neonatal data). Six studies did not contribute any data to the review

Study Reference	**Brown 2017I** (Insulin SLR)			
			Metformin arm range: 22 (29–39) (n=35) to 36 (IQR 35,37) (n=14) *Insulin vs glibenclamide (N=11)* Insulin arm range: 26 ± 3.4 (n=30) to 32.6 ± 6.2 (n=46) Glibenclamide arm range: 24.9 ± 3.7 (n=10) to 32.2 ± 5.0 (n=41) *Insulin vs acarbose (N=2)* Insulin arm: 28.7 ± 6.0 (n=27) (NR for other study) Acarbose arm: 31.5 ± 5.8 (n=19) (NR for other study) *Insulin vs glyburide/metformin combined (N=3)* Insulin arm: 30.7 ± 4.4 (n=33) to 32.1 ± 5.7 (n=42) (NR for 1 study) Glyburide/metformin arm: 31.1 ± 4.7 (n=35) to 33.2 ± 4.9 (n=42) (NR for 1 study) **Insulin type A vs type B** *Human insulin vs insulin aspart (N=5)* Human insulin arm: 29.6 ± 4.5 (n=157) to 31.0 ± 2.7 (n=5) Insulin aspart arm: 29.2 ± 4.0 (n=163) to 31.6 ± 5.9 (n=14) *Human insulin vs insulin lispro (N=3)* Human insulin arm: 29.8 ± 1.0 (n=23) to 35 (28–41) (n=24) (NR for 1 study) Insulin lispro arm: 34.2 ± 1.3 (n=19) to 34.5 (28–41) (n=25) (NR for 1 study) *Human insulin vs neutral protamine Hagedorn insulin (N=2)* Human insulin arm: NR Neutral protamine Hagedorn insulin arm: NR *Insulin detemir vs neutral protamine Hagedorn insulin (N=1)* Insulin detemir arm: 35 (IQR 31–38) (n=42) Neutral protamine Hagedorn insulin arm: 35 (IQR 32–38) (n=45) **Insulin vs diet** (N=4) Insulin arm: 27 ± 5.4 to 31.8 (n=47) Diet arm: 26 ± 5.7 to 32.7 (n=56) **Insulin vs exercise** (N=1) Insulin arm: NR Exercise arm: NR **Insulin vs standard care** (N=2) Insulin arm: NR Standard care arm: NR **Insulin regimen A vs insulin regimen B** (N=2) Regimen A arm: 33 ± 5 (n=136) (NR for 1 study)	

Study Reference	**Brown 2017l** (Insulin SLR)			
			Regimen B arm: 33 ± 5 (n=138) (NR for 1 study)	

Study Reference	**Brown 2017l** (Insulin SLR)

		Ethnicity/race	NR: N=25 NR but likely to be same as setting: N=15 Ethnicity reported (reported ethnicities include Mexican, White, Black, Hispanic, Native American, Alaskan, Caucasian, African American, Jewish, Bantu (Zulu), Polynesian, Indian): N=13

Study Reference	**Brown 2017l** (Insulin SLR)		
		Maternal BMI at baseline (kg/m^2)	**Insulin vs oral antidiabetic** *Insulin vs metformin (N=19)* Insulin arm range: 25.1 ± 3.4 (n=50) to 37.9 ± 6.87 (n=14) Metformin arm range: 26.4 ± 2.8 (n=50) to 39.5 ± 6.94 (n=16) *Insulin vs glibenclamide (N=11)* Insulin arm range: 22.6 ± 3.1 (n=129) to 31.8 ± 5.1 (n=59) Glibenclamide arm range: 21.9 ± 2.80 (n=120) to 33.4 ± 12.9 (n=41) *Insulin vs acarbose (N=2)* Insulin arm: 27.0 ± 7.2 (n=27) to 32.1 ± 5.6 (n=46) Acarbose arm: 25.7 ± 4.2 (n=19) to 33.1 ± 6.4 (n=45) *Insulin vs glyburide/metformin combined (N=3)* Insulin arm: 32.2 ± 7.2 (n=33) (NR for 2 studies) Glyburide/metformin arm: 32.0 ± 5.4 (n=35) (NR for 2 studies) **Insulin type A vs type B** *Human insulin vs insulin aspart (N=5)* Human insulin arm: 25.6 ± 2.9 (n=5) to 33.2 ± 5.7 (n=13) Insulin aspart arm: 26.0 ± 3.4 (n=163) to 29.3 ± 4.7 (n=14) *Human insulin vs insulin lispro (N=3)* Human insulin arm: 22.3 (19.8_25.3) (n=24) to 33.3 ± 1.2 (n=23) Insulin lispro arm: 21.5 (19.2_25.1) (n=25) to 31.5 ± 1.1 (n=19) *Human insulin vs neutral protamine Hagedorn insulin (N=2)* Human insulin arm: NR Neutral protamine Hagedorn insulin arm: NR *Insulin detemir vs neutral protamine Hagedorn insulin (N=1)* Insulin detemir arm: 28.3 (IQR 24.9_33.8) (n=42) Neutral protamine Hagedorn insulin arm: 28.6 (IQR 24.4_31.1) (n=45) **Insulin vs diet** (N=4) Insulin arm: NR Diet arm: NR **Insulin vs exercise** (N=1) Insulin arm: NR Exercise arm: NR **Insulin vs standard care** (N=2) Insulin arm: NR Standard care arm: NR **Insulin regimen A vs insulin regimen B** (N=2) Regimen A arm: 27.8 ± 2.7 (n=136) (NR for 1 study) Regimen B arm: 27.9 ± 2.6 (n=138) (NR for 1 study)	

Study Reference **Brown 2017I** (Insulin SLR)

Study Reference	**Brown 2017l** (Insulin SLR)			
		Gestational age at intervention start (weeks)	**Insulin vs oral antidiabetic** *Insulin vs metformin (N=19)* Insulin arm range: 26.7 ± 3.5 (n=50) to 32.05 ± 3.50 (n=47) Metformin arm range: 27.6 ± 3.3 (n=50) to 32.18 ± 3.70 (n=47) *Insulin vs glibenclamide (N=11)* Insulin arm range: 22.6 ± 5.6 (n=13) to 30.6 ± 2.2 (n=41) Glibenclamide arm range: 22.5 ± 4.7 (n=10) to 30.8 ± 2.5 (n=41) *Insulin vs acarbose (N=2)* Insulin arm: 30.2 ± 3.7 (n=46) (NR for 1 study) Acarbose arm: 30.5 ± 3.5 (n=45) (NR for 1 study) *Insulin vs glyburide/metformin combined (N=3)* Insulin arm: 24.5 ± 6.3 n=42) to 30.1 ± 3.1 (n=33) (NR for 1 study) Glyburide/metformin arm: 22.1 ± 7.3 (n=42) to 29.3 ± 3.8 (n=35) (NR for 1 study) **Insulin type A vs type B** *Human insulin vs insulin aspart (N=5)* Human insulin arm: 22.4 ± 10.1 (n=157) (NR for 4 studies) Insulin aspart arm: 21.7 ± 9.3 (n=163) (NR for 4 studies) *Human insulin vs insulin lispro (N=3)* Human insulin arm: 25.6 ± 1.3 (n=23) to 29 (27–32) (n=24) (NR for 1 study) Insulin lispro arm: 27.3 ± 1.4 (n=19) to 29 (26–32) (n=25) (NR for 1 study) *Human insulin vs neutral protamine Hagedorn insulin (N=2)* Human insulin arm: NR Neutral protamine Hagedorn insulin arm: NR *Insulin detemir vs neutral protamine Hagedorn insulin (N=1)* Insulin detemir arm: 27.3 (IQR 23.3–28.5) (n=42) Neutral protamine Hagedorn insulin arm: 28.1 (IQR 25.1–29.3) (n=45) **Insulin vs diet** (N=4) Insulin arm: NR Diet arm: NR **Insulin vs exercise** (N=1) Insulin arm: NR Exercise arm: NR **Insulin vs standard care** (N=2) Insulin arm: NR Standard care arm: NR	

Study Reference	**Brown 2017l** (Insulin SLR)			
			Insulin regimen A vs insulin regimen B (N=2) Regimen A arm: 28 ± 6.9 (n=136) (NR for 1 study) Regimen B arm: 27.4 ± 6.8 (n=138) (NR for 1 study)	

Study Reference	**Brown 2017l** (Insulin SLR)	
	Diagnostic criteria for GDM	Carpenter and Coustan 1983 (N=12) Carpenter and Coustan 1983 or IADPSG 2010 (N=1) WHO 1994 (N=5) ADA (N=5); ADA 2012 (N=1); ADA 2011 (N=1); earlier ADA criteria (N=3) Australian Diabetes in Pregnancy Society 1998 (N=3) National Diabetes Data Group 1979 (N=2) Canadian Diabetes Association (no date specified) (N=1) IADPSG 2010 (N=2) Modified O'Sullivan and Mahan (N=1) Gillmer 1975 (N=1) Finnish National Guidelines 2008 (N=2) NR (N=18)
	Treatment targets	
	Interventions and comparisons	**Insulin vs oral antidiabetic** Insulin vs metformin (N=19) – 4 did not contribute data to the review Insulin vs glibenclamide (N=11) Insulin vs acarbose (N=2) Insulin vs combined metformin and glibenclamide (N=3) **Insulin type A vs type B** (N=10) **Insulin vs diet** (N=4) **Insulin vs exercise** (N=1) **Regimens of insulin** (N=2)
Study Reference	(Insulin SLR) **Maternal primary outcomes** Hypertensive disorders of pregnancy (any definition) (N=13) Caesarean section (N=25) Development of type 2 diabetes (N=4) **Neonatal primary outcomes** Perinatal (foetal and neonatal death) (N=15) LGA (N=19) Death or serious morbidity composite (N=3) Neurosensory disability in later childhood (N=1) **Maternal secondary outcomes**	
	Outcomes	Use of additional pharmacotherapy (N=23); Maternal hypoglycaemia (N=16); Glycaemic control during/end of treatment (fasting) (N=25); Glycaemic control during/end of treatment (postprandial) (N=23); Glycaemic control during/end of treatment (HbA1c) (N=15); Weight gain in pregnancy (N=14) Induction of labour (N=3); Postpartum haemorrhage (N=2); Breastfeeding at discharge, 6 weeks postpartum, 6 months or longer

Brown 2017I	(N=2); Maternal mortality (N=1); BMI (N=1); Impaired glucose tolerance (N=3); Biomarker changes (N=1) **Neonatal secondary outcomes** Stillbirth (N=3); Neonatal death (N=3); Macrosomia (N=29); SGA (N=13); Birth trauma (N=16); Gestational age at birth (N=24); Preterm birth (N=15); Congenital abnormality (N=17); 5-minut Apgar score <7 (N=5); Birthweight (N=33); Head circumference at birth (N=3); Length at birth (N=7); Ponderal index at birth (N=5); Adiposity (N=2); Neonatal hypoglycaemia (N=31); Respiratory distress syndrome (N=13); Neonatal jaundice (hyperbilirubinaemia) (N=21); Hypocalcaemia (N=7); Polycythaemia (N=5); Biomarker changes (N=4); Childhood weight (N=2); Childhood height (N=2); Childhood adiposity (N=2);

Study Reference (Insulin SLR)

	Childhood blood pressure (N=1); Number of antenatal visits or admissions (N=1); Admission to NICU/nursery (N=18); Duration of stay in NICU (N=3)
Funding	Academic/government funding not related to pharmaceutical industry (N=10) Statement of no funding received (N=4) Funding related to pharmaceutical industry (N=3) No statement about funding (N=36)
Conflicts of interest	Statement of no conflicts of interest (N=14) No details on conflicts of interest (N=36) Conflict of interest reported for 1 or more author (N=3)

	Brown 2017I
Methods	Searches Sources searched: • Pregnancy and Childbirth's Trials Register, ClinicalTrials.gov, WHO International Clinical Trials Registry Platform (ICTRP), reference lists of retrieved studies ○ The Cochrane Pregnancy and Childbirth's Trials Register is maintained by an Information Specialist and contains trials identified from CENTRAL, MEDLINE, Embase, CINAHL (EBSCO), hand-searches of 30 journals and the proceedings of major conferences, weekly current awareness alerts for a further 44 journals plus monthly BioMed Central email alerts Screening and selection process Eligibility, risk of bias and data extraction were performed by two review authors. Data were checked for accuracy. Any disagreements were resolved by a discussion or a third person was consulted if necessary Study quality assessment Risk of bias was independently assessed by two review authors using criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011). Disagreement was resolved by discussion or by involving a third assessor Methods for combining intervention evidence • Dichotomous data: summary risk ratio with 95% CIs • Continuous data: mean difference if outcomes were measured in the same way between trials; standardised mean difference to combine trials that measure the same outcome, but used different methods • For all outcomes, analyses were carried out on ITT basis (as far as possible) • Heterogeneity in each meta-analysis was assessed using Tau^2, I^2 and Chi^2. Heterogeneity was regarded as substantial if I^2 >30% and either Tau^2>0 or P<0.10 in the Chi^2 test • If there were ≥10 studies in the meta-analysis, reporting biases (such as publication bias) were assessed using funnel plots (asymmetry assessed visually) • Fixed-effects meta-analysis was used where it was reasonable to assume that studies were estimating the same underlying treatment effect (sufficiently similar population and methods etc.) • Random-effects meta-analysis was used if there was heterogeneity – also investigated using subgroup and sensitivity analyses

Comparison 1: insulin vs oral antidiabetic agent

Study Reference **Brown 2017I** (Insulin SLR)

Adverse maternal and neonatal outcomes

Maternal outcomes

Patient or population: the treatment of women (maternal outcomes) with GDM

Setting: primary and secondary care (Canada, Egypt, USA, Brazil, Finland, Iran, Australia, New Zealand, India)

Intervention: insulin

Comparison: oral antidiabetic pharmacological therapy

Outcome, n (%)	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	N of participants (studies)	Quality of evidence (GRADE)	Comments
	Oral antidiabetic agent	Insulin				
Hypertensive disorders of pregnancy (pre-eclampsia)	77 per 1000	88 per 1000 (66–117)	RR 1.14 (0.86–1.52)	2060 (10 RCTs)	Moderate[1]	No data were reported for eclampsia
Hypertensive disorders of pregnancy (no defined)	36 per 1000	69 per 1000 (42–114)	RR 1.89 (1.14–3.12)	1214 (4 RCTs)	Moderate[1]	There were no definitions for hypertensive disorders of pregnancy in the trials reporting this outcome
Caesarean section	394 per 1000	405 per 1000 (366–449)	RR 1.03 (0.93–1.14)	1988 (17 RCTs)	Moderate[1]	
Development of type 2 diabetes	52 per 1000	73 per 1000 (42–128)	RR 1.39 (0.80–2.44)	754 (2 RCTs)	Moderate[2]	These 2 trials compared insulin with metformin. No other trials reported this long-term outcome
Perineal trauma/tearing	This was not measured by any trial					
Postnatal weight retention or return to pre-pregnancy weight (maternal weight 6–8 weeks postpartum)	Mean weight: 80.8 kg	Mean difference: –1.6 kg (–6.34–+3.14)	Mean difference: –1.6 kg (–6.34–+3.14)	167 (1 RCT)	Low[2,3]	
Postnatal weight retention or return to pre-pregnancy weight	Mean weight: 81.8 kg	Mean difference: –3.7 kg (–8.5–+1.1)	Mean difference: –3.7 kg (–8.5–+1.1)	176 (1 RCT)	Low[2,3]	

Study Reference **Brown 2017I** (Insulin SLR)

	(maternal weight 1 year postpartum)						
	Postnatal depression	This was not measured by any trial					
	Induction of labour	408 per 1000	535 per 1000 (424–669)	Average RR 1.30 (0.96–1.75)	348 (3 RCTs)	Moderate[2]	These 3 trials compared insulin with metformin. No other trials reported this outcome

* The risk in the intervention group (and 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI)

[1] Risk of bias: most of the trials were not blinded – downgraded one level

[2] Risk of bias: no blinding, lacked methodological details to be able to judge randomisation or allocation concealment – downgraded one level [3] Imprecision: wide confidence intervals and single study – downgraded one level

Study Reference **Brown 2017l** (Insulin SLR)

Neonatal outcomes
Patient or population: infants of women with GDM
Setting: primary and secondary care (Canada, Egypt, USA, Brazil, Finland, Iran, Australia, New Zealand, India)
Intervention: insulin
Comparison: oral antidiabetic pharmacological therapy

Outcome, n (%)	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	N of participants (studies)	Quality of evidence (GRADE)	Comments
	Oral antidiabetic agent	Insulin				
LGA (birthweight >90th centile)	159 per 1000	161 per 1000 (121–215)	Average RR 1.01 (0.76–1.35)	2352 (13 RCTs)	Moderate[1]	
Perinatal and later infant mortality	8 per 1000	7 per 1000 (2–20)	RR 0.85 (0.29–2.49)	1463 (10 RCTs)	Low[1,2]	Event rates were low (6/735 for antidiabetic pharmacological therapies group and 5/728 for insulin group). No data were reported for later infant mortality
Death or serious morbidity composite	319 per 1000	329 per 1000 (268–402)	RR 1.03 (0.84–1.26)	760 (2 RCTs)	Moderate[1]	These 2 trials compared insulin with metformin. No other trials reported this outcome. One trial included resuscitation in the delivery room, preterm birth <37 weeks, NICU admission, birth injury or diagnosis of neonatal complication, glucose infusion, antibiotics or phototherapy. The other trial included hypoglycaemia <2.6 mmol/L, respiratory distress syndrome, phototherapy, birth trauma, 5-minute Apgar <7, preterm birth <37 weeks

Study Reference **Brown 2017I** (Insulin SLR)

Neonatal hypoglycaemia	111 per 1000	126 per 1000 (94– 169)	Average RR 1.14 (0.85–1.52)	3892 (24 RCTs)	Low[1,5]	
Adiposity at birth – % fat mass	Mean 12.8	Mean difference –1.6 (– 3.77–+0.57)	Mean difference –1.60 (– 3.77–+0.57)	82 (1 RCT)	Moderate[4]	
Adiposity at birth – skinfold thickness (mm)	Mean 16	Mean difference –0.8 (–0.49–+0.73)	Mean difference –0.8 (–2.33–+0.73)	82 (1 RCT)	Very low[2,4,7]	
Adiposity in childhood – % fat mass	Mean 16.4	Mean difference +0.5 (– 0.49–+1.49)	Mean difference +0.5 (– 0.49–+1.49)	318 (1 RCT)	Low[1,4]	

Study Reference **Brown 2017I** (Insulin SLR)

Child/adulthood diabetes (type 1/2)	This was not measured by any trial					
Neurosensory disability at 18 months – mild developmental delay	104 per 1000	111 per 1000 (34–358)	RR 1.07 (0.33–3.44)	93 (1 RCT)	Low[4,6]	
Neurosensory disability at 18 months – hearing impairment	0 per 1000	0 per 1000 (0–0)	RR 0.31 (0.01–7.49)	93 (1 RCT)	Low[4,6]	
Neurosensory disability at 18 months – visual impairment	21 per 1000	6 per 1000 (1–60)	RR 0.31 (0.03–2.90)	93 (1 RCT)	Low[4,6]	

* The risk in the intervention group (and 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI)

[1] Risk of bias: most of the trials were not blinded – downgraded one level

[2] Imprecision: event rates are low and CIs are wide cross the line of no effect – downgraded one level

[3] Inconsistency: I^2 = 78% – downgraded one level

[4] Evidence based on single trial – downgraded one level

[5] Inconsistency: I^2 = 51% – downgraded one level

[6] Imprecision: wide CIs – downgraded one level

[7] Risk of bias: selective reporting and other bias detected – downgraded one level

Risk of bias summary for each study

Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Anjalakshi 2007	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	HIGH	HIGH	HIGH
Ardilouze 2014	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	HIGH	HIGH	LOW
Ashoush 2016	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	LOW	LOW
Balaji 2005	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	HIGH	HIGH
Balaji 2012	UNCLEAR	LOW	HIGH	UNCLEAR	1. LOW	HIGH	LOW
Behrashi 2016	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	UNCLEAR	LOW

Study Reference **Brown 2017l** (Insulin SLR)

Bertini 2005	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	LOW	LOW
Beyou 2015	LOW	LOW	HIGH	LOW	UNCLEAR	UNCLEAR	LOW
Bung 1993	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	HIGH	HIGH
Castorino 2011	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	UNCLEAR
Coustan 1978	HIGH	HIGH	HIGH	UNCLEAR	LOW	HIGH	HIGH
De Veciana 2002	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	HIGH	HIGH	HIGH
Di Cianni 2007	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	HIGH	LOW
Hague 2003	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	HIGH	UNCLEAR
Herrera 2015	LOW	LOW	HIGH	HIGH	UNCLEAR	HIGH	UNCLEAR
Hickman 2013	LOW	LOW	HIGH	UNCLEAR	HIGH	HIGH	HIGH
Hutchinson 2008	LOW	LOW	HIGH	UNCLEAR	UNCLEAR	HIGH	HIGH
Ijas 2011	LOW	LOW	HIGH	UNCLEAR	LOW	UNCLEAR	LOW
Ismail 2007	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	HIGH
Jovanovic 1999	LOW	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	UNCLEAR
Lain 2009	LOW	LOW	UNCLEAR	LOW	UNCLEAR	HIGH	HIGH
Langer 2000	LOW	LOW	HIGH	HIGH	LOW	UNCLEAR	LOW
Majeed 2015	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	LOW
Martinez Piccole 2010	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	HIGH	HIGH
Mecacci 2003	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW
Mesdaghinia 2013	LOW	LOW	LOW	LOW	LOW	UNCLEAR	LOW
Mirzamoradi 2015	LOW	LOW	UNCLEAR	UNCLEAR	LOW	UNCLEAR	UNCLEAR
Mohamed 2014	LOW	LOW	UNCLEAR	LOW	UNCLEAR	UNCLEAR	HIGH
Moore 2007	LOW	LOW	HIGH	UNCLEAR	LOW	UNCLEAR	HIGH
Mukhopadhyay 2012	LOW	UNCLEAR	HIGH	HIGH	LOW	HIGH	LOW
Nachum 1999	LOW	LOW	UNCLEAR	UNCLEAR	LOW	HIGH	LOW
Niromanesh 2012	LOW	LOW	HIGH	UNCLEAR	LOW	LOW	LOW
Notelovitz 1971	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	HIGH	HIGH
O'Sullivan 1975a	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	UNCLEAR
O'Sullivan 1975b	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	HIGH	HIGH
Ogunyemi 2007	LOW	2. LOW	HIGH	UNCLEAR	LOW	HIGH	HIGH
Pavithra 2016	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	LOW
Persson 1985	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	HIGH	HIGH	LOW

Study Reference **Brown 2017I** (Insulin SLR)

Pettitt 2007	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	HIGH	LOW
Poyhonen-Alho 2002	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	HIGH
Prasad 2008	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	HIGH	HIGH	HIGH
Riaz 2014	LOW	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	HIGH
Rowan 2008	LOW	LOW	HIGH	UNCLEAR	LOW	UNCLEAR	UNCLEAR
Ruholamin 2014	LOW	LOW	LOW	LOW	LOW	UNCLEAR	LOW
Saleh 2016	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	LOW
Silva 2007	UNCLEAR	UNCLEAR	HIGH	HIGH	LOW	UNCLEAR	HIGH
Spaulonci 2013	LOW	UNCLEAR	UNCLEAR	UNCLEAR	LOW	HIGH	LOW
Tertti 2013	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	LOW	LOW	LOW
Thompson 1990	LOW	UNCLEAR	HIGH	UNCLEAR	LOW	HIGH	LOW
Waheed 2013	LOW	UNCLEAR	HIGH	UNCLEAR	LOW	UNCLEAR	LOW
Wali 2015	UNCLEAR	LOW	HIGH	UNCLEAR	UNCLEAR	HIGH	HIGH
Zangeneh 2014	UNCLEAR	UNCLEAR	HIGH	UNCLEAR	UNCLEAR	LOW	LOW
Zawiejska 2016	LOW	LOW	HIGH	UNCLEAR	LOW	HIGH	LOW

Authors' Conclusions

- Main comparison in the review was insulin vs oral antidiabetic pharmacological therapies. Insulin and oral antidiabetic pharmacological therapies have similar effects on key health outcomes. The quality of evidence ranged from very low to moderate, with downgrading decisions due to imprecision, risk of bias and inconsistency
- For the other comparisons (insulin vs non-pharmacological interventions; different insulin analogues; different insulin regimens), there is insufficient volume of high-quality evidence to determine differences for key health outcomes
- Long-term maternal and neonatal outcomes were poorly reported for all comparisons
- The evidence suggests that there are minimal harms associated with the effects of treatment with either insulin or oral antidiabetic pharmacological therapies. The choice to Use one or the other may be down to physician or maternal preference, availability or severity of GDM. Further research is needed to explore optimal insulin regimens. Further research could aim to report data for standardised GDM outcomes

Abbreviations: GDM, gestational diabetes mellitus; NR, not reported; RCT, randomised controlled trial

Study Reference **Brown 2017I** (Insulin SLR)

Table 94: Brown 2017A (Antidiabetics)

Study Reference	**Brown 2017A (Antidiabetic SLR)**
Study Design	Design Systematic literature review Objective To evaluate the effects of oral anti-diabetic pharmacological therapies for treating women with GDM. Search dates 14th May 2016 Country Various Setting NR

Study Reference **Brown 2017A (Antidiabetic SLR)**

Study eligibility

Inclusion (PICOS)

Population	Pregnant women diagnosed with gestational diabetes (diagnosis as defined by the individual trial). Women with known type 1 or type 2 diabetes were excluded
Intervention	Antidiabetic pharmacological therapies used during pregnancy, including metformin, glibenclamide, acarbose, tolbutamide, chlorpropamide, or any combination of these therapies. Only oral antidiabetic therapies were included.
Comparator	Standard care or another lifestyle intervention
Outcomes	**Primary** *Maternal* • Hypertensive disorders of pregnancy (including preeclampsia, pregnancy-induced hypertension, eclampsia as defined by trialists • Caesarean section • Development of type 2 diabetes *Neonatal* • Perinatal (fetal and neonatal death) and later infant mortality • Large-for-gestational age (LGA) (as defined by trialists) • Death or serious morbidity composite (variously defined by trials, e.g. perinatal or infant death, shoulder dystocia, bone fracture or nerve palsy) • Neurosensory disability in later childhood (as defined by trialists) **Secondary** *Maternal* Use of additional pharmacotherapy; maternal hypoglycaemia (as defined by trialists); glycaemic control during/end of treatment (as defined by trialists); weight gain in pregnancy; adherence to the intervention; induction of labour; placental abruption; postpartum haemorrhage (as defined by trialists; postpartum infection; perineal trauma/tear; breastfeeding at discharge, six weeks postpartum, six months or longer; maternal mortality; sense of well-being and quality of life; behavioural changes associated with the intervention; views of the intervention; relevant biomarker changes associated with the intervention (including adiponectin, free fatty acids, triglycerides, high-density lipoproteins (HDL), low-density lipoproteins (LDL), insulin) *Long-term outcomes for mother* Postnatal depression; body mass index (BMI); postnatal weight retention or return to pre-pregnancy weight; type 1 diabetes; type 2 diabetes; impaired glucose tolerance; subsequent gestational diabetes; cardiovascular health (as defined by trialists including blood pressure, hypertension, cardiovascular disease, metabolic syndrome) *Fetal/neonatal outcomes*

Study Reference **Brown 2017A (Antidiabetic SLR)**

	Stillbirth; neonatal death; macrosomia (greater than 4000 g; or as defined by individual study); small-for-gestational (SGA) age (as defined by trialists); birth trauma (shoulder dystocia, bone fracture, nerve palsy); gestational age at birth; preterm birth (< 37 weeks' gestation; and < 32 weeks' gestation); five-minute Apgar less than seven; birthweight and z score; head circumference and z score; length and z score; ponderal index; adiposity (including skinfold thickness measurements (mm); fat mass as defined by trialists); neonatal hypoglycaemia (as defined by trialists); respiratory distress syndrome; neonatal jaundice (hyperbilirubinaemia) (as defined by trialists); hypocalcaemia (as defined by trialists); polycythaemia (as defined by trialists); relevant biomarker changes associated with the intervention (including insulin, cord c-peptide) *Later infant/childhood outcomes* Weight and z scores; height and z scores; head circumference and z scores; adiposity (including BMI, skinfold thickness, fat mass); educational attainment; blood pressure; type 1 diabetes; type 2 diabetes; impaired glucose intolerance; dyslipidaemia or metabolic syndrome *Child as an adult outcomes* Weight; height; adiposity (including BMI, skinfold thickness, fat mass); employment, education and social status/achievement; dyslipidaemia or metabolic syndrome; type 1 diabetes; type 2 diabetes; impaired glucose tolerance; cardiovascular health (as defined by trialists including blood pressure, hypertension, cardiovascular disease, metabolic syndrome) *Health service use* Number of antenatal visits or admissions; number of hospital or health professional visits (including midwife, obstetrician, physician, dietician, diabetic nurse); admission to neonatal intensive care unit/nursery; duration of stay in neonatal intensive care unit or special care baby unit; length of antenatal stay; length of postnatal stay (maternal); length of postnatal stay (baby); cost of maternal care; cost of offspring care; costs associated with the intervention; costs to families associated with the management provided; cost of dietary monitoring (e.g. diet journals, dietician, nurse visits, etc); costs to families - change of diet, extra antenatal visits; extra use of healthcare services (consultations, blood glucose monitoring, length and number of antenatal visits); women's view of treatment advice; duration of stay in neonatal intensive care unit or special care baby unit; duration of material and neonatal hospital stay (antenatal, neonatal, postnatal)
Study design	Published or unpublished randomised controlled trials in full-text or abstract format

Exclusion (reasons given in excluded study list)

- Women with type 1 or type 2 diabetes
- Non-RCTs
- Cross-over trials
- Postpartum intervention
- Non-interventional studies

Study Reference **Brown 2017A (Antidiabetic SLR)**

Other NR

Flow of Studies (PRISMA)

- Database results: 253
- Hand-searches/other sources: 2
- 254 records after duplicates removed
- Title/abstracts reviewed: 254
- Full-texts reviewed: 28
- Articles included in qualitative synthesis: 11 from 20 publications
  - Articles included in quantitative synthesis (meta-analysis): 8 studies

Included study characteristics

Characteristic	Details
Design	All 11 included studies were RCTs with a parallel design; (20 publications, of which 3 published as conference abstracts)
Sample sizes	Minimum 40 – maximum 207 participants Total of 1487 women and their babies
Setting and timing	Brazil (N=3); USA (N=3); India (N=2); South Africa (N=1); UK (N=1); Israel (N=1)
Participants	Maternal age was reported in 6/11 trials, maternal BMI in 5/11 trials, and gestational age at study entry was reported in 4/11 trials

Study	Maternal age, mean (SD)	Maternal BMI (kg/m^2)	Gestational age at study entry (weeks), mean (SD)
Bertini 2005	Glibenclamide: 31.2 (4.5); Acarbose: 31.5 (5.8)	Glibenclamide: 27.5 (5.8); Acarbose: 25.7 (4.2)	NR
Casey 2015	Glibenclamide: 31.3 (6.0); Placebo: 31.2 (6.0)	Glibenclamide: 29.0 (4.8); Placebo: 28.9 (5.3)	Glibenclamide: 26.0 (2.0); Placebo: 26.0 (1.0)
Cortez 2006	NR	NR	NR
De Bacco 2015	NR	NR	NR
Fenn 2015	NR	NR	NR
George 2015	Metformin: 33.4 (4.4); Glibenclamide: 33.6 (4.6)	Metformin: 28.7 (4.4); Glibenclamide: 28.8 (4.0)	Metformin: 29.3 (3.3); Glibenclamide: 29.7 (SD 3.7)
Moore 2010	Metformin: 31 (7.1); Glibenclamide: 29.6 (7.8)	Metformin: 32.8 (5.8); Glibenclamide: 32.7 (7.0)	Metformin: 27.3 (6.8); Glibenclamide: 29.1 (5.0)
Myers 2014	NR	NR	NR
Nachum 2015	NR	NR	NR
Notelovitz 1971	Cholpropramide: 30.9; Tolbutamide: 29.7	NR	NR

Study Reference	Brown 2017A (Antidiabetic SLR)	

	Silva 2012	Metformin: 32.6 (5.6) Glibenclamide: 31.3 (5.4)	Metformin: 28.7 (5.4) Glibenclamide: 28.6 (5.9)	Metformin: 27.0 (6.4) Glibenclamide 25.4 (7.1)

Diagnostic criteria for GDM	Four different diagnostic criteria, where details were provided • Carpenter and Coustan criteria: 3 trials (Fenn 2015, Moore 2010, Nachum 2015) • National Diabetes Data Group (NDGG 1979): 2 trials (George 2015 and Casey 2015) • World Health Organization (1999): 2 trials (Bertini 2005 and Silvia 2012) • Unspecified WHO criteria: 1 trial (De Bacco 2015) • 3 trials did not provide details on the criteria used to diagnose the women with gestational diabetes (Cortez 2006, Myers 2014, Notelovitz 1971)
Treatment targets	Eight studies reported 3 different treatment targets for fasting blood glucose: • less than 5.0 mmol/L (90 mg/dL) (Bertini 2005; Silva 2012); • less than 5.3 mmol/L (95 mg/dL) (Casey 2015; Cortez 2006; Fenn 2015; George 2015; Nachum 2015); • less than 5.8 mmol/L (105 mg/dL) (Moore 2010). Four studies reported four different treatment targets for one-hour postprandial blood glucose: • less than 6.7 mmol/L (120 mg/dL) (Silva 2012); • less than 7.2 mmol/L (130 mg/dL) (George 2015) 90 minutes; • less than 7.5 mmol/L (135 mg/dL) (Cortez 2006); • less than 7.8 mmol/L (140 mg/dL) (Fenn 2015). Four studies reported two different treatment targets for two-hour postprandial blood glucose: • less than 5.5 mmol/L (100 mg/dL) (Bertini 2005); • less than 6.7 mmol/L (120 mg/dL) (Casey 2015; George 2015; Moore 2010) Notelovitz 1971 reported “good control” was obtained with a postprandial (timing not specified) blood glucose reading of less than 8.3 mmol/L (150 mg/dL). De Bacco 2015 and Myers 2014 did not provide details of glycaemic targets for treatment. The Notelovitz 1971 study included women with both GDM and pre-gestational diabetes. The proportions of women included with each type of diabetes were unclear and therefore the data have not been included in any meta-analyses.

Study Reference	Brown 2017A (Antidiabetic SLR)
Interventions and comparisons	Four different comparisons were reported Oral anti-diabetic pharmacological therapy vs placebo or usual care • Glibenclamide vs placebo: 1 trial (Casey 2015) • Acarbose vs placebo: 1 trial (Cortez 2006) • Metformin vs standard care: 1 trial (Myers 2014) • Tolbutamide with chlorpropamide vs diet (Notelobitz 1979) Metformin vs glibenclamide: 5 trials (De Bacco 2015, Fenn 2015, George 2015, Moore 2010, Silva 2012)
	Glibenclamide vs acarbose: 1 trial (Bertini 2005) Glibenclamide with or without metformin vs metformin with or without glibenclamide: 1 trial (Nachum 2015; glibenclamide plus metformin if glycaemic targets were not met, with metformin plus glibenclamide if glycaemic targets were not met)
Outcomes	
Funding	Notelovitz 1971: declared financial support received from Pfizer laboratories. No other studies provided details on funding sources.
Conflicts of interest	Fenn 2015 and Moore 2010: authors declared no conflict of interest. No other studies provided details on conflicts of interest.

Definition of GDM

Diagnosis as defined by the individual trial.

Study Reference	Brown 2017A (Antidiabetic SLR)
Methods	Searches Sources searched: Cochrane Pregnancy and Childbirth's Trials Register, which is maintained by an Information Specialist and contains trials identified from: • CENTRAL (monthly searches) • MEDLINE (weekly searches via Ovid) • Embase (weekly searches via Ovid) • CINAHL (monthly searches via EBSCO) • Hand-searches of 30 journals and the proceedings of major conferences • Weekly current awareness alerts for a further 44 journals plus monthly BioMed Central email alerts Screening and selection process Search results were screened by two people and the full text of all relevant trial reports identified through the searches were reviewed. Based on the intervention described, each trial report was assigned a number that corresponds to a specific Pregnancy and Childbirth review topic (or topics), and was then added to the register. Two review authors independently assessed for inclusion all the potential studies identified as a result of the search strategy. Any disagreement was resolved through discussion, or if required, through consultation with a third person. Study quality assessment. Two reviewers independently assessed risk of bias for each randomised study using the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions. Any disagreement was resolved by discussion or by involving a third assessor. Quality assessment The quality of the evidence was assessed using the GRADE approach. Up to a maximum of seven outcomes were selected for the mother and seven for the infant, covering both short- and long-term outcomes for the main comparisons Methods for combining intervention evidence

Study Reference
Brown 2017A (Antidiabetic SLR)

A fixed-effect meta-analysis was used for combining data where it was reasonable to assume that studies were estimating the same underlying treatment effect: i.e. where trials examined the same intervention, and the trials' populations and methods were judged sufficiently similar. If clinical heterogeneity was sufficient to expect that the underlying treatment effects differed between trials, or if substantial statistical heterogeneity was detected, a random-effects meta-analysis was produced for an overall summary, if an average treatment effect across trials was considered clinically meaningful. The random-effects summary was treated as the average of the range of possible treatment effects and we discussed the clinical implications of treatment effects differing between trials. If the average treatment effect was not clinically meaningful, we did not combine trials. Where we used random-effects analyses, the results were presented as the average treatment effect with 95%confidence intervals, and the estimates of Tau^2 and I^2.

Glibenclamide vs placebo

Study Reference **Brown 2017A (Antidiabetic SLR)**

Outcome	Anticipated absolute effects[a] (95% CI)		Risk ratio (95% CI)	Participants N (studies)	Quality of evidence (GRADE)	Comments
	Risk with placebo	Risk with glibenclamide				
Pregnancy outcomes						
Gestational age at birth	NR	NR	NR	NR	NR	
Hypertensive disorders of pregnancy (pre-eclampsia)	per 1000	per 1000 (135–317)	1.24 (0.81–1.90)	(1 RCT)	Very low	
Perinatal mortality (fetal and neonatal death) and later infant mortality	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome
Mode of birth	NR	NR	NR	NR	NR	
Induction of labour	per 1000	per 1000 (149–331)	1.18 (0.79–1.76)	(1 RCT)	Very low	
Vaginal delivery	NR	NR	NR	NR	NR	
Instrumental delivery	NR	NR	NR	NR	NR	
C-section (not specified if planned or emergency)	per 1000	per 1000 (285–483)	1.03 (0.79–1.34)	(1 RCT)	Very low	
Maternal gestational weight gain	NR	NR	NR	NR	NR	
Preterm delivery	NR	NR	NR	NR	NR	
Perineal trauma/tear	per 1000	per 1000 (0–84)	0.98 (0.06–15.62)	(1 RCT)	Very low	
Maternal outcomes						
Maternal wellbeing	NR	NR	NR	NR	NR	
Postpartum haemorrhage	NR	NR	NR	NR	NR	
Method of infant feeding	NR	NR	NR	NR	NR	
Postnatal depression	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome

Study Reference	Brown 2017A (Antidiabetic SLR)						
Effectiveness of the Intervention	*Postnatal weight retention or return to pre-pregnancy weight*	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome

Study Reference **Brown 2017A (Antidiabetic SLR)**

Post-pregnancy T2DM	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome
Neonatal outcomes						
Macrosomia	NR	NR	NR	NR	NR	
LGA	118 per 1000	105 per 1000 (60–187)	0.89 (0.51–1.58)	375 (1 RCT)	Very low	
Composite outcome (death, shoulder dystocia, nerve palsy, bone fracture)	NR	NR	NR	NR	NR	
Adiposity (neonatal) – neonatal fat mass (g)	NR	NR	NR	NR	NR	
Shoulder dystocia	NR	NR	NR	NR	NR	
Brachial plexus neuropathy	NR	NR	NR	NR	NR	
Neonatal hypoglycaemia	11 per 1000	21 per 1000 (4–114)	1.97 (0.36–10.62)	375 (1 RCT)	Very low	Event rates were low with 4/189 for oral antidiabetic pharmacological therapy and 2/186 for placebo group
Admission to NICU	NR	NR	NR	NR	NR	

GRADE Working Group grades of evidence: High quality: We are very conf ident that the true effect lies close to that of the estimate of the effect; **Moderate quality:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different; **Low quality:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect; **Very low quality:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Metformin vs glibenclamide

Outcome	Anticipated absolute effects[a] (95% CI)		Risk ratio (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)	Comments
	Risk with glibenclamide	Risk with metformin				
Pregnancy outcomes						
Gestational age at birth	NR	NR	NR	NR	NR	-
Hypertensive disorders of pregnancy (pre-eclampsia)	88 per 1000	62 per 1000 (33–114)	0.70 (0.38–1.30)	508 (2 RCTs)	Moderate	-

Study Reference **Brown 2017A (Antidiabetic SLR)**

Perinatal mortality (fetal and neonatal death) and later infant mortality	6 per 1000	5 per 1000 (0– 83)	0.92 (0.06–14.55)	359 (2 RCTs)	Very low	Note that event rates were very low. 1 study had no event of perinatal death in either the metformin or the glibenclamide
						group. The second study had 1 death in each group
Mode of birth	NR	NR	NR	NR	NR	-
Induction of labour	613 per 1000	496 per 1000 (374–655)	0.81 (0.61–1.07)	159 (1 RCT)	Low	
Vaginal delivery	NR	NR	NR	NR	NR	-
Instrumental delivery	NR	NR	NR	NR	NR	-
C-section (not specified if planned or emergency)	392 per 1000	470 per 1000 (325–674)	1.20 (0.83–1.72)	554 (4 RCTs)	Low	-
Maternal gestational weight gain	NR	NR	NR	NR	NR	-
Preterm delivery	NR	NR	NR	NR	NR	-
Perineal trauma/tear	6 per 1000	11 per 1000 (1–81)	1.67 (0.22–12.52)	308 (2 RCTs)	Low	Note low event rates (2/154 for metformin and 1/154 for glibenclamide)
Maternal outcomes						
Maternal wellbeing	NR	NR	NR	NR	NR	-
Postpartum haemorrhage	NR	NR	NR	NR	NR	-
Method of infant feeding	NR	NR	NR	NR	NR	-
Postnatal depression	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome
Postnatal weight retention or return to pre-pregnancy weight	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome
Post-pregnancy T2DM	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome
Neonatal outcomes						-
Macrosomia	NR	NR	NR	NR	NR	-

Study Reference **Brown 2017A (Antidiabetic SLR)**

LGA	193 per 1000	129 per 1000 (46–354)	0.67 (0.24–1.83)	246 (2 RCTs)	Low	-
Death or serious comorbidity composite	350 per 1000	189 per 1000 (109–329)	0.54 (0.31–0.94)	159 (1 RCT)	Low	-
Adiposity (neonatal) – neonatal fat mass (g)	See comment	See comment	NR	NR	NR	None of the included studies prespecified this outcome
Shoulder dystocia	NR	NR	NR	NR	NR	-
Brachial plexus neuropathy	NR	NR	NR	NR	NR	-
Neonatal hypoglycaemia	48 per 1000	41 per 1000 (20–84)	0.86 (0.42–1.77)	554 (4 RCTs)	Low	-
Admission to NICU	NR	NR	NR	NR	NR	-

GRADE Working Group grades of evidence: High quality: We are very conf ident that the true effect lies close to that of the estimate of the effect; **Moderate quality:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a

Study Reference	Brown 2017A (Antidiabetic SLR)

possibility that it is substantially different; **Low quality:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect; **Very low quality:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Study Reference **Brown 2017A (Antidiabetic SLR)**

Glibenclamide vs acarbose

Outcome	Anticipated absolute effects[a] (95% CI)		Risk ratio (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)	Comments
	Risk with placebo	Risk with glibenclamide				
Pregnancy outcomes						
Gestational age at birth	NR	NR	NR	NR	NR	-
Hypertensive disorders of pregnancy (pre-eclampsia)	NR	NR	NR	NR	NR	-
Perinatal mortality (fetal and neonatal death) and later infant mortality	NR	NR	NR	NR	NR	-
Mode of birth	NR	NR	NR	NR	NR	-
Induction of labour	NR	NR	NR	NR	NR	-
Vaginal delivery	NR	NR	NR	NR	NR	-
Instrumental delivery	NR	NR	NR	NR	NR	-
C-section (not specified if planned or emergency)	526 per 1000	500 per 1000 (279–895)	0.95 (0.53–1.70)	43 (1 RCT)	Low	-
Maternal gestational weight gain	NR	NR	NR	NR	NR	-
Preterm delivery	NR	NR	NR	NR	NR	-
Perineal trauma/tear	NR	NR	NR	NR	NR	-
Maternal outcomes						-
Maternal wellbeing	NR	NR	NR	NR	NR	-
Postpartum haemorrhage	NR	NR	NR	NR	NR	-
Method of infant feeding	NR	NR	NR	NR	NR	-
Postnatal depression	NR	NR	NR	NR	NR	-
Postnatal weight retention or return to pre-pregnancy weight	NR	NR	NR	NR	NR	-
Post-pregnancy T2DM	NR	NR	NR	NR	NR	-
Neonatal outcomes						
Macrosomia	NR	NR	NR	NR	NR	-
LGA	105 per 1000	251 per 1000 (57–1000)	2.38 (0.54–10.46)	43 (1 RCT)	Low	-

Death or serious comorbidity composite	NR	NR	NR	NR	NR	-
Adiposity (neonatal) – neonatal fat mass (g)	NR	NR	NR	NR	NR	-

Shoulder dystocia	NR	NR	NR	NR	NR	-
Brachial plexus neuropathy	NR	NR	NR	NR	NR	-
Neonatal hypoglycaemia	53 per 1000	333 per 1000 (46–1000)	6.33 (0.87–46.32)	43 (1 RCT)	Low	Low event rates and sample size (8/24 in glibenclamide group; 1/19 in acarbose group)
Admission to NICU	NR	NR	NR	NR	NR	-

GRADE Working Group grades of evidence: High quality: We are very conf ident that the true effect lies close to that of the estimate of the effect; **Moderate quality:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different; **Low quality:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect; **Very low quality:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Abbreviations: CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio.

Risk of bias summary: review authors' judgements about each risk of bias item for each included

	) Random sequence generator selection bias	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	) Blinding of outcome assessment detection bias	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias

Study Reference	**Brown 2017A (Antidiabetic SLR)**								
Study Reference	**Brown 2017A (Antidiabetic SLR)**								
	[redacted]								
	Bertini 2005	❑	❑	❑	❑	❑	❑	❑	
	Casey 2015	❑	❑	❑	❑	❑	❑	❑	
	Cortez 2006	❑	❑	❑	❑	❑	❑	❑	
	De Bacco 2015	❑	❑	❑	❑	❑	❑	❑	
	Fenn 2015	❑	❑	❑	❑	❑	❑	❑	
	George 2015	❑	❑	❑	❑	❑	❑	❑	
	Moore 2010	❑	❑	❑	❑	❑	❑	❑	
	Myers 2014	❑	❑	❑	❑	❑	❑	❑	
	Nachum 2015	❑	❑	❑	❑	❑	❑	❑	
	Notelovitz 1971	❑	❑	❑	❑	❑	❑	❑	
	Silva 2012	❑	❑	❑	❑	❑	❑	❑	

Authors' Conclusions

There were insufficient data comparing oral anti-diabetic pharmacological therapies with placebo/standard care (lifestyle advice) to inform clinical practice. There was insufficient high-quality evidence to be able to draw any meaningful conclusions as to the benefits of one oral anti-diabetic pharmacological therapy over another due to limited reporting of data for the primary and secondary outcomes in this review. Short- and long-term clinical outcomes for this review were inadequately reported or not reported. Current choice of oral anti-diabetic pharmacological therapy appears to be based on clinical preference, availability and national clinical practice guidelines. The benefits and potential harms of one oral anti-diabetic pharmacological therapy compared with another, or compared with placebo/ standard care remains unclear and requires further research. Future trials should attempt to report on the core

Study Reference **Brown 2017A (Antidiabetic SLR)** outcomes suggested in this review, in particular long-term outcomes for the woman and the infant that have been poorly reported to date, women's experiences and cost benefit.

Appendix 4 – Guidance on quality assessments

Guidance used

Table 95. Guidance on the use of PROBAST

Question	Literature-recommended criteria GDM specific
TYPE of PREDICTION STUDY	
Classify the evaluation based on its aim (i.e. what is the type of prediction study)? (Development only/Development and validation/Validation only)	Development only if there is prediction model development without external validation. These studies may include internal validation methods e.g. bootstrapping and cross-validation techniques Development and validation if there is prediction model development combined with external validation in other participants in the same article Validation only if external validation of an existing (previously developed) model in other participants
PARTICIPANTS	

Risk of Bias	1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data? (Y/PY/PN/N/NI)	Higher potential for RoB when participant data are from existing sources (e.g. existing cohort studies or routine care registries) because their data are often collected for a different purpose than a model Study design with a lowest RoB for a diagnostic model is a cross sectional study where a group (cohort) of participants is selected on the basis of certain symptoms or signs that make them suspected of having the target condition of interest. Randomised intervention trials can also be used, however the randomised treatments may need to be included as separate predictors to account for any treatment effects. In addition, the inclusion criteria in RCTs are usually more restricted, resulting in narrower "predictor distributions". Models developed/validated using data with narrower predictor distributions tend to show lower discriminative ability than those with more broadly distributed predictors For case-cohort or nested case-control studies, low RoB can be considered so long as authors appropriately adjust for the original cohort/registry outcome frequency in the analysis (also applies to question 4.6 later). If not, they are at high RoB because they are from an 'existing source' (i.e. sampled from another cohort or registry). There is further guidance for prognostic models	Y/PY if cohort design (including RCT or proper registry data) or a nested case-control/-cohort with adjustment for baseline risk/hazard in the analysis N/PN if a non-nested case-control (or any other study design)
	1.2 Were all inclusions and exclusions of participants appropriate? (Y/PY/PN/N/NI)	This question relates to inclusion/exclusion at the at the enrolment stage (e.g. not loss-to-follow-up). The key issue is whether any inclusion or exclusion criteria or recruitment strategy could have made the included study participants unrepresentative of the target population Example: inappropriate inclusion results from including participants already known to have the outcome at the time of the predictor measurement; this will most likely result in a model with overestimated predictive performance	Y/PY if inclusion/exclusion appropriate i.e. participants reflect unselected participants of interest N/PN if included participants would already have been identified as having the outcome or if specific subgroups excluded that may have altered the performance of the predictive model for the intended target population
	What is the risk of bias introduced by selection of participants (Low/High/Unclear)	Low if answer to all signalling questions is Y or PY. If ≥1 of the answers is N or PN, the judgement could still be low but specific reasons should be provided as to why it can be considered so High if the answer to any signalling questions is N or PN, unless otherwise defined as low above Unclear if relevant information is missing for some of the signalling questions and none were judged at high RoB	
Applicability	What is the concern that the included participants and setting do not match the review question? (Low/High/Unclear)	Included participants, the selection criteria used and the setting used in the primary prediction model study should be relevant to the review question Low if included participants and clinical setting match the review question High if included participants and clinical setting differ from the review question Unclear if relevant information is not reported	
PREDICTORS			

Risk of Bias	2.1 Were predictors defined and assessed in a similar way for all participants? (Y/PY/PN/N/NI)	Potential for this bias is higher for predictors that involve subjective judgement e.g. imaging test results (risk of looking at predictive ability of observer rather than predictor)	Y/PY if definitions of predictors and their assessment were similar for all participants N/PN if different definitions were used for the same predictor or if predictors requiring subjective interpretation were assessed by differently experience assessors
	2.2 Were predictor assessments made without knowledge of outcome data? (Y/PY/PN/N/NI)	I.e. blinding or masking. This is also especially important for predictors that involve subjective interpretation or judgement Blinding of assessors to outcome naturally occurs in prognostic studies with a prospective cohort design where the predictors are assessed before the outcome has happened. Bias is more likely in studies that retrospectively record predictors (recall bias) or if predictors and outcomes are assessed at a similar time (crosssectional studies) If no information on blinding is given, this domain can still be rated as low RoB in overall assessment if predictors were measured/reported a long time before the outcome If predictors are collected by reinterpreting stored data (i.e. samples), assessors may be aware of the outcome	Y/PY if outcome information was stated as not used during predictor assessment or was clearly not (yet) to those assessing predictors N/PN if it is clear that outcome information was used when assessing predictors
	2.3 Are all predictors available at the time the model is intended to be used? (Y/PY/PN/N/NI)	I.e. would they be available when the model is intended to be used on a patient at the time of prediction Studies that aim to externally validate existing prediction models are at high RoB when predictor data are missing at the time of validation and the authors validate the model anyway by omitting the missing predictors. This is a common flaw in validation studies (i.e. validating a different model than the original). In these cases, this signalling question should be answered N.	Y/PY if all included predictors would be available at the time the model is intended to be used for prediction N/PN if predictors would not be available at the time the model is intended to be used for prediction
	What is the risk of bias introduced by predictors or their assessment? (Low/High/Unclear)	Low if answer to all signalling questions is Y or PY. If ≥1 of the answers is N or PN, the judgement could still be low but specific reasons should be provided as to why it can be considered so e.g. use of objective predictors not requiring subjective interpretation High if the answer to any signalling questions is N or PN, unless otherwise defined as low above Unclear if relevant information is missing for some of the signalling questions and none were judged at high RoB	
Applicability	What is the concern that the definition, assessment or timing of predictors in the model do not match the review question? (Low/High/Unclear)	Low if the definition, assessment, and timing of predictors match the review question High if the definition, assessment, or timing of predictors were different from the review question Unclear if relevant information about the predictors is not reported	
OUTCOME			
Risk of Bias	3.1 Was the outcome determined	This is about the level of measurement error within the method of determining the outcome (see concerns for applicability about whether the definition of the outcome is appropriate)	Y/PY if a method of outcome determination has been used which is considered optimal or acceptable by guidelines or previous publications on the topic
	appropriately? (Y/PY/PN/N/NI)	If prediction model study uses data from routine care registries or existing studies originally designed/conducted to answer a different research question, their outcome determination methods should be appraised Potential for bias is higher in outcomes that involve subjective judgement, such as imaging, surgical or pathology results	N/PN if a clearly suboptimal method that causes unacceptable error in determining outcome status has been used

	3.2 Was a prespecified or standard outcome definition used? (Y/PY/PN/N/NI)	RoB is low when a prespecified/standard outcome definition is used and substantiated by a definition from clinical guidelines/previously published study/study protocol RoB is higher if, e.g., an atypical threshold on a continuous scale has been used to defined the outcome – this may be evident if authors test multiple thresholds to obtain the most favourable outcome definition Composite outcomes can also introduce RoB e.g. if model performance is adjusted by excluding typical components and excluding atypical components Many outcomes have consensus-based definitions. Determining whether standard or non-standard definitions have been used may require specialist clinical knowledge	Y/PY if the method of outcome determination is objective or if a standard definition is used or if prespecified categories are used to group outcomes N/PN if the outcome definition was not standard and not prespecified
	3.3 Were predictors excluded from the outcome definition? (Y/PY/PN/N/NI)	In some cases, it is not possible to avoid including predictors in outcome determination. E.g., if the outcome is decided by a consensus panel using as much information as is available. If a model predictor forms part of the definition or assessment of the outcome, the association between predictor and outcome will likely by overestimated (incorporation bias)	Y/PY if none of the predictors were included in the outcome definition N/PN if ≥1 of the predictors forms part of the outcome definition
	3.4 Was the outcome defined and determined in a similar way for all participants? (Y/PY/PN/N/NI)	E.g., same thresholds and categories; same method of combining individual components if a composite outcome; same method for establishing the outcome in consensus- or panel-based decisions (e.g. majority vote) Look out for variation between research sites in multicentre studies RoB is higher in models that are based on data collected for a different purpose (e.g. registry, existing study) as inherently different outcome definitions are likely to be applied If outcome is dependent on accuracy of measurement or subjective interpretation, along with if outcomes are measured on several occasions at different frequency for different participants (more frequent visits = more likely to detect), RoB is higher	Y/PY if outcomes were defined and determined in a similar way for all participants N/PN if outcomes were clearly defined/determined in a different way for some participants
	3.5 Was the outcome determined without knowledge of predictor information? (Y/PY/PN/N/NI)	Similar to 3.3 In consensus or panel decisions on outcome, it may be that as much information as possible is available, which could include the predictor If the aim of a model is to assess the incremental value of a certain predictor or compare the performance of competing models (i.e. validating >1 model on the same data set), the importance of blinded outcome determination is higher	Y/PY if predictor information was not known when determining the outcome status N/PN if it is clear that predictor information was used when determining the outcome status
	3.6 Was the time interval between predictor assessment and outcome determination appropriate? (Y/PY/PN/N/NI)	Bias can present in two ways: 1. Outcome determined too early, when relevant outcome cannot be detected or the number of outcomes is unrepresentative 2. Type of outcome may differ depending on time interval, e.g. metastases detected early may be liver metastases, whereas at one year they may mainly be bone metastases Time interval is also relevant to applicability of the review and whether you are trying to determine short- or long-term prognosis	Y/PY if the time interval between predictor assessment and outcome determination was appropriate to enable the correct type and representative number of relevant outcomes to be recorded or if no information on time interval is needed to enable this N/PN if the time interval is too long or too short to enable the correct type and representative number of relevant outcomes to be recorded

	What is the risk of bias introduced by the outcome or its determination? (Low/High/Unclear)	Low if answer to all signalling questions is Y or PY. If ≥1 of the answers is N or PN, the judgement could still be low but specific reasons should be provided as to why it can be considered so e.g. if outcome was determined with knowledge of predictor information but the outcome assessment did not require much interpretation by the assessor e.g. death from any cause High if the answer to any signalling questions is N or PN, unless otherwise defined as low above Unclear if relevant information is missing for some of the signalling questions and none were judged at high RoB	
Applicability	What is the concern that the outcome, its definition, timing or determination do not match the review question? (Low/High/Unclear)	Low if outcome definition, timing, and method of determination defines the outcome as intended by the review question. High if choice of outcome definition, timing, and method of outcome determination defines another outcome as intended by the review question. Unclear if relevant information about the outcome, timing, and method of determination is not reported.	
ANALYSIS			
Risk of Bias	4.1 Were there a reasonable number of participants with the outcome? (Y/PY/PN/N/NI)	Model development studies Performance of any prediction model is overestimate (to some extent) when development and assessment of performance both use the same data set – overestimation is larger with smaller sample size and when fewer participants have the outcome, and when model predictors are selected from a large number of candidate predictors (i.e. those considered during the model development process) EPV (events per variable) = number of participants with the outcome relative to the number of candidate predictor parameters *For EPV between 10–20, the item should be rated as PY or PN, depending on the outcome frequency, model performance and distribution of predictors in the model The lower the EPV, the higher the likelihood that the model has been 'overfitted' or 'underfitted' (included spurious predictors or failed to include important predictors)Consider if the predictors used in the model would be typically assessed in woman being screened for GDM Model validation studies Because the aim in a validation study is accurate and precise estimation of model performance, they are recommended to include at least 100 participants	Model development studies Y/PY if EPV ≥20* N/PN if EPV <10* Model validation studies Y/PY if number of participants with the outcome is ≥100 N/PN if number of participants with the outcome is <100

	4.2 Were continuous and categorical predictors handled appropriately? (Y/PY/PN/N/NI)	Both Dichotomisation of continuous variables (predictors) requires choosing (often) an arbitrary cut-off point, which leads to loss of information and reduced predictive ability of the model (e.g. two people may have very different values but both be above the cutoff so would be classified as the same) This is particularly a problem if the cut-off is chosen to maximise the predictive effect of the model Model development studies Low RoB when predictors are kept continuous. The association between predictor and outcome risk should still be examined as linear or nonlinear RoB can still be low if a model categorises continuous predictors into 4 or more groups, rather than dichotomises, especially if these are based on widely accepted cut-offs. However, it should be clear that cut-offs were chosen before the data analysis Model validation studies Predictors should have the same format in the model validation study as they did in the development	Both Y/PY if continuous predictors are not converted into ≥2 categories (dichotomised) when included in the model or if continuous predictors are examined for nonlinearity or if categorical predictor groups are defined using a prespecified method N/PN if categorical predictor groups do not use a prespecified method Model development studies Y/PY *No extra criteria* N/PN if continuous predictors are converted into ≥2 categories when included in the model Model validation studies Y/PY if continuous predictors use the same definitions/transformations and categorical predictors are categorised using the same cut points as in the development study N/PN if they use different definitions

	4.3 Were all enrolled participants included in the analysis? (Y/PY/PN/N/NI)	For lowest RoB, all enrolled patients should be included. If low %s are excluded from the analysis, RoB may still be low, but 'low' % is hard to define because it depends on which participants were excluded and whether this was a selected subsample or not Model studies based on existing sources (existing study or care database/registry) are particularly susceptible to this type of bias. In such cases, participant selection for the analysis should be based on clear criteria	Y/PY if all participants enrolled in the study are included in the analysis N/PN if some or a subgroup of participants are inappropriately excluded from the analysis
	4.4 Were participants with missing data handled appropriately? (Y/PY/PN/N/NI)	When a study report does not mention missing data, participants with any missing data have likely been omitted from the analyses ("available-case" or "complete-case" analysis) because statistical packages automatically exclude persons with any missing value on any of the data analysed unless prompted otherwise The most appropriate method for handling missing data is multiple imputation because it leads to the least biased results, whilst missing indicator method (using a separate category to capture missing data) leads to biased results If authors provide further details (e.g. comparison of with- and without missing values), a more informed judgement on the RoB can be made (i.e. if there is not much difference, RoB may still be low) If a model validation study is using data where a specific predictor is missing (e.g. because it was not measured), simply omitting the predictor leads to high RoB and this question should be rated as N	Y/PY if there are no missing values of predictors or outcomes and the study explicitly reports that participants are not excluded on the basis of missing data or if missing values are handled using multiple imputation N/PN if participants with missing data are omitted from the analysis or if the method of handling missing data is clearly flawed (e.g. missing indicator method or inappropriate use of last value carried forward) or if the study had no explicit mention of methods to handle missing data

	4.5 Was selection of predictors based on univariable analysis avoided? (Y/PY/PN/N/NI)	In a univariable analysis, individual predictors are tested for their association with the outcome and those with a statistically significant univariable association are often selected for inclusion in the development of the model. This can lead to incorrect predictor selection because they are chosen on the basis of their significance as a single predictor rather than in combination with other predictors This can lead to bias if some predictors are omitted that should not be – some predictors are only important after adjustment for others. Predictors may also be selected by accidental association with the outcome using this approach A better approach is to use non-statistical methods, e.g., existing knowledge of established predictors Some statistical methods that are not based on prior statistical tests between predictor and outcome can be used to reduce the number of modelled predictors (e.g. principal component analysis)	Y/PY if the predictors are <u>not</u> selected on the basis of univariable analysis prior to multivariable modelling N/PN if the predictors <u>are</u> selected on the basis of univariable analysis prior to multivariable modelling
	4.6 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately? (Y/PY/PN/N/NI)	For case-cohort/case-controls, the analysis method must account for the sampling fractions (from the original cohort) For prognostic models to predict long-term outcomes where censoring occurs, a time-to-event analysis (e.g. Cox regression) should be used to include censored participants up to the end of their follow-up. Excluding censored patients with incomplete follow-up is inappropriate. Competing risks should also be appropriately accounted for If a person can have >1 event, multilevel or random effects modelling methods are needed to avoid underestimation	Y/PY if complexities in the data are accounted for appropriately <u>or</u> if they have been identified appropriately as unimportant N/PN if data complexities that could affect model performance are ignored
	4.7 Were relevant model performance measures	Model calibration and discrimination should be assessed appropriately	Y/PY if both calibration and discrimination are evaluated appropriately (including relevant measures tailored for models predicting survival outcomes)

	evaluated appropriately? (Y/PY/PN/N/NI)	Calibration: agreement between predictions from model and observed outcomes, preferably reported graphically (calibration plot). Calibration is frequently assessed by calculating the Hosmer-Lemeshow goodness-of-fit test; however this has limited suitability to evaluate poor calibration Discrimination: ability of model to distinguish between individuals who do or do not develop the outcome. The most widely reported measure of discrimination is the concordance index (c-index), which is equivalent to the area under the receiver-operating characteristic (ROC) curve for logistic regression models Calibration and discrimination measures should account for the type of outcome being predicted. For survival models, researchers should account for time-to-event and censoring using e.g. Harrell's c-index or the D statistic Classification measures such as sensitivity, specificity or predictive value may also be used. These require the introduction of one or more threshold in the range of model-predicted probabilities which allows reporting of the model's performance at probability thresholds which may be clinically relevant. However, use of thresholds leads to loss of information and choice of thresholds may be data-driven rather than prespecified, which can lead to bias (i.e. thresholds chosen to maximise performance). The choice of threshold should be prespecified for low RoB	N/PN if both calibration and discrimination are not evaluated or if only goodness-of-fit tests (e.g. Hosmer-Lemeshow test) are used to evaluate calibration or if for models predicting survival outcomes performance measures accounting for censoring are not used or if classification measures (sensitivity, specificity, predictive values) were presented using predicted probability thresholds derived from the data set at hand
	4.8 Were model overfitting and optimism in model performance accounted for? (Y/PY/PN/N/NI)	This applies to model development studies only Studies developing models should always include some form of internal validation (i.e. using data of the original sample) e.g. bootstrapping and cross-validation If optimism is present, an important next step is to adjust or shrink the model predictive performance estimates and predictor effects, however this is not typically done and will lead to bias The need to adjust for overfitting and optimism is greater for studies with a small sample size and low EPV and those using stepwise predictor selection strategies	Y/PY if internal validation techniques, such as bootstrapping and crossvalidation including all model development procedures, have been used to account for any optimism in model fitting, and subsequent adjustment of the model performance (e.g. shrinkage) estimates have been applied N/PN if no internal validation has been performed or if internal validation consists only of a single random split-sample of participant data or if the bootstrapping or cross-validation did not include all model development procedures (including any variable selection)
	4.9 Do predictors and their assigned weights in the final model correspond to the results from the reported multivariable analysis? (Y/PY/PN/N/NI)	This applies to model development studies only Predictors and coefficients for the developed model, including intercept or baseline components, should be fully reported to allow others to correctly apply the model to other individuals The final presented model and the results from the multivariable analysis should match, otherwise bias may arise.	Y/PY if predictors and regression coefficients in the final model correspond to reported results from multivariable analysis N/PN if predictors and regression coefficients in the final model do not correspond to reported results from the multivariable analysis
	What is the risk of bias introduced by the analysis? (Low/High/Unclear)	Low if answer to all signalling questions is Y or PY. If ≥1 of the answers is N or PN, the judgement could still be low but specific reasons should be provided as to why it can be considered so High if the answer to any signalling questions is N or PN, unless otherwise defined as low above Unclear if relevant information is missing for some of the signalling questions and none were judged at high RoB	
OVERALL ASSESSMENT			

Overall risk of bias judgement (Low risk of bias/High risk of bias/Unclear risk of bias)	Low risk of bias if all domains were rated at low risk of bias. For models developed without any external validation, only consider at low RoB if all domains rated as low and the model's development was based on a very large data set and included some form of *internal* validation – otherwise, consider high risk of bias High risk of bias if at least one domain is judged to be at high risk of bias Unclear risk of bias if unclear risk of bias was noted in at least one domain and it was low risk for all others
Overall applicability judgement (Low concerns for applicability/High concerns for applicability/Unclear concerns for applicability)	Low concerns for applicability if it is judged as such for all domains High concerns for applicability if it is judged as such for at least one domain Unclear concerns for applicability if it is judged as unclear for at least one domain and there are no domains judged as high concerns for applicability

Table 96. Guidance on the use of Cochrane ROB

Question	Response Options	Literature-Recommended Criteria
RANDOMISATION PROCESS	Yes (Y), Possibly Yes (PY), Possibly No (PN), No information (NI), Not applicable (NA)	
1.1 Was the allocation sequence random?	Y, PY, PN, N, NI	Y if random component was used in sequence generation process (e.g. computer-generated random numbers, random number table, coin tossing). Use of minimization technique can also be considered random
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y, PY, PN, N, NI	Y if the process of allocation is controlled by an external unit or organisation, independent of the enrolment personnel (e.g. telephone or internet-based) If envelopes or drug containers used, adequate detail should be given e.g. to the level that envelopes are opaque, sequentially numbered, sealed with a tamper-proof seal and irreversibly assigned to the participant. If this detail is not provided, should assign PY or PN N if reason to suspect that investigator or participant was aware of the allocation N if no random element used (e.g. alternation; methods based on dates [of birth or admission]; patient record numbers; allocation decisions made by clinicians or participants; allocation based on the availability of the intervention; or any other systematic or haphazard method] PY if judged likely to be random e.g. experienced clinical trials unit with absence of specific information about generation of randomised sequence in paper with tight word limit
1.3 Did the baseline differences between intervention groups suggest a problem with the randomization process?	Y, PY, PN, N, NI	N if no apparent imbalances or if imbalances are likely due to chance Y if there are imbalances that indicate problems with the randomisation e.g. large difference in intervention group size, imbalance in ≥1 key prognostic baseline characteristics, or conversely, if baseline characteristics are excessively similar NI if there is no useful baseline information available PN if e.g. other trials by same investigator/team have used non-random approaches
Risk of bias judgement	Low, High, Some concerns	As per the algorithm at https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool?authuser=0
EFFECT OF ASSIGNMENT TO INTERVENTION		
2.1 Were participants aware of their assigned during the trial?	Y, PY, PN, N, NI	N if trial was blinded, however, if participants experience side effects or toxicities that could be attributed to one of the interventions, the answer should be Y or PY
2.2 Were carers and people delivering the assigned intervention during the trial?	Y, PY, PN, N, NI	N if trial was blinded, however, if participants experience side effects or toxicities that could be attributed to one of the interventions, the answer should be Y or PY If randomisation allocation was not concealed, it is likely that carers/people delivering intervention were aware of the assignment

2.3 If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?	NA, Y, PY, PN, N, NI	The term 'trial context' refers to the effects of recruitment/engagement activities on trial participants e.g. seeking informed consent (so a patient knows their allocation) may lead patients in a placebo group to seek other intervention Y or PY only if there is evidence that the trial context led to failure to implement the protocol or starting of interventions not allowed by the protocol N or PN if there were changes from the protocol, but these could occur outside of the trial context e.g. non-adherence to an intervention
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA, Y, PY, PN, N, NI	Deviations will only impact the intervention effect estimate if they affect the outcome
2.5 If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA, Y, PY, PN, N, NI	Deviations are more likely to impact the intervention effect estimate if they are not balanced between groups

Question	Response Options	Literature-Recommended Criteria
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y, PY, PN, N, NI	ITT and mITT (excluding participants with missing outcome data) analyses should be considered appropriate Per protocol and as treated analyses should be considered inappropriate Analyses excluding eligible patients post-randomisation are inappropriate, but excluding ineligible patients post-randomisation (e.g. if eligibility was not yet confirmed) are appropriate
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA, Y, PY, PN, N, NI	There is no precise rule. It is possible that even if <5% of participants were analysed in the wrong group or excluded, this could have a substantial impact on the results
Risk-of-bias judgement	Low, High, Some concerns	As per the algorithm at https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool?authuser=0

MISSING OUTCOME DATA

3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y, PY, PN, N, NI	'Nearly all' = the number of participants with missing outcome data is sufficiently small that their outcomes would have made no important difference to the estimated effect of the intervention For continuous outcomes, availability of data for 95% of the participants will often be sufficient. For dichotomous outcomes, the proportion required is directly linked to the risk of the event – if the observed number of events is much greater than the number of missing data, the bias will be small Only report NI if no information is given about missing outcome data – this will usually lead to a judgement that there is a high risk of bias Imputed data should be regarded as missing data for this question
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?	NA, Y, PY, PN, N	Y or PY if there are analysis methods that correct for bias or sensitivity analyses showing that results are little changed under a range of assumptions about the relationship between missing outcomes and its true value Imputation (e.g. 'last-observation-carried-forward') should not be assumed to correct for bias due to missing outcome data
3.3 If N/PN to 3.2: Could missingness in the true value?	NA, Y, PY, PN, N, NI	N/PN if missing outcome data occurred for reasons unrelated to the outcome, the risk of bias due to this outcome will be low Y/PY if it was related to the participant's health status (i.e. discontinuation of study due to adverse effects) In time-to-event analyses, participants censored from the analysis should be considered as having missing data
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA, Y, PY, PN, N, NI	Possible reasons for answering Y are: differences between intervention groups in terms of amount of missing outcome data; reported reasons for missing outcome data suggest that it depends on the true value or differ between intervention groups; in time-to-event analyses, if follow-up is censored when participants stop or change their intervention e.g. due to toxicity or a need for second-line chemotherapy
Risk-of-bias judgement	Low, High, Some concerns	See algorithm on crib sheet

MEASUREMENT OF OUTCOME

4.1 Was the method of measuring the outcome inappropriate?	Y, PY, PN, N, NI	In most cases, for pre-specified outcomes, the answer will be N or PN Y or PY if the method of data collection is inappropriate e.g. it is unlikely to be sensitive to intervention effects (e.g. ranges of outcome values are not detectable using the method) or the measurement instrument has been shown to have poor validity
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	Y, PY, PN, N, NI	N or PN if data collection involves the same measurement methods and thresholds (including number of times measures are taken) across intervention arms
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?	NA, Y, PY, PN, N, NI	N if outcome assessors were blinded to the intervention status. For patient-reported outcomes, the patient should be blinded
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA, Y, PY, PN, N, NI	Outcomes that are likely influenced by knowledge of the intervention are ones which involve some level of judgement (e.g. level of pain), rather than e.g. all-cause mortality
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA, Y, PY, PN, N, NI	If there are strong levels of belief in either harmful or beneficial effects of the intervention, it is more likely that the outcome was influenced by knowledge

Question	Response Options	Literature-Recommended Criteria
Risk-of-bias judgement	Low, High, Some concerns	As per the algorithm at https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool?authuser=0
SELECTION OF THE REPORTED RESULT		
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y, PY, PN, N, NI	If available, planned outcome measurements/analyses can be compared with those presented in published reports. Finalisation of analysis plans must be before unblinded data become available to investigators Changes to analysis plans made before unblinded outcome data were available (or unrelated to the results) do not raise concerns for bias
5.2. Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	Y, PY, PN, N, NI	It may be possible to report certain outcomes in more than one way (e.g. for pain, different scales, at different timepoints). If this is done but results are only reported for one particular method, there is a high risk of bias in the fully reported result Y or PY if there is clear evidence that a domain was measured in multiple eligible ways but data for only a subset of measures is reported (without justification) and the selection was likely influenced by the result of that subset (e.g. more significant) N or PN if there is only one way an outcome can be measured or if results for all eligible measures are reported
5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible analyses of the data?	Y, PY, PN, N, NI	It may be possible to analyse outcomes in more than one way (e.g. adjusted and unadjusted models, absolute value and change from baseline). As above, if multiple estimates are generated but only one subset reported, there is a high risk of bias Y or PY if there is clear evidence that outcomes were analysed in multiple eligible ways but data for only a subset of analyses is reported (without justification) and the selection of this reporting was likely influenced by its result N or PN if there is only one way the outcome could be analysed or if results for all analyses conducted are reported
Risk-of-bias judgement	Low, High, Some concerns	As per the algorithm at https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool?authuser=0
OVERALL RISK OF BIAS	Low, some concerns, high	Low if the study is judged to be at low RoB for all domains Some concerns if the study is judged to raise some concerns in at least one domain, but is not at high RoB for any domain High if the study is judged to be at high RoB in at least one domain or the study is judged to have some concerns for multiple domains in a way that substantially lowers confidence in the result

Table 97. Guidance on the use of ROBINS-I

Question	Response options	Literature-recommended criteria	Criteria for GDM Rapid Review
BIAS DUE TO CONFOUNDING			
1.1 Is there potential for confounding of the effect of intervention in this study?	Y/PY/PN/N/NI (Yes/Probably yes/Probably no/No/No Information)	Factors likely to influence the effect of interventions at baseline, for example, if age influences the appearance of hyperglycaemia (and therefore placement of the woman in a "hyperglycaemic" vs "normal" category) this is a source of confounding bias	Factors likely to influence the effect of interventions at baseline: maternal age, ethnicity, BMI, week of gestation at diagnosis with GDM, use of other medications or prophylaxis received prior to glucose measurement
1.2 If Y/PY to 1.1: Was the analysis based on splitting participants' follow up time according to intervention received?	Y/PY/PN/N/NI	If participants could switch between intervention groups then associations between intervention and outcome may be biased by time-varying confounding. This occurs when prognostic factors influence switches between intended interventions.	
1.4 Did the authors use an appropriate analysis method that controlled for all the important confounding domains? If there were analyses to control for (assuming the	Not applicable(NA)/Y/PY/PN/N/NI) confounders include stratification, regression, matching, standardization, and inverse probability	Appropriate methods to control for measured ethnicity, BMI were controlled for (or the uniform across the group) weighting. They may control for individual	Consider at low risk if maternal age, confounding variables,

	same variables as listed in 1.1) answer Each method depends on the assumption that there were differences between groups in residual confounding.	variables or for the estimated propensity score. should be these variables at baseline.	Y or PY, depending on whether there is no unmeasured or
1.5 If Y/PY to 1.4 Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	Y/PY/PN/N/NI	Appropriate control of confounding requires that the variables adjusted for are valid and reliable measures of the confounding domains. Study authors may cite references to support the use of a particular measure. If authors control for confounding variables with no indication of their validity or reliability pay attention to the subjectivity of the measure. Subjective measures (e.g. based on self-report) may have lower validity and reliability than objective measures such as lab findings.	Age, ethnicity, BMI: likely to be measured reliably, so only consider answering the question with N or PN if the authors controlled for other variables that may not have been measured reliably
1.6 Did the authors control for any post-intervention?	NA/Y/PY/PN/N/NI Have the authors controlled for variables that are measured after the intervention is	intervention variables that could have been affected by the	

Response options	Literature-recommended criteria	Criteria for GDM Rapid Review	
	received? Are these variables likely to be affected by intervention or affect the outcome of the intervention?		
Risk of bias judgement	Low/ Moderate/ Serious/ Critical/ NI **Low**: no confounding expected **Moderate**: (i) Confounding expected, all known important confounding domains appropriately measured and controlled for; **and** (ii) Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding. **Serious**: (i) At least one known important domain was not appropriately measured, or not controlled for; **or** (ii) Reliability or validity of measurement of an important domain was low enough that we expect serious residual confounding. **Critical**: (i) Confounding inherently not controllable, **or** (ii) The use of negative controls strongly suggests unmeasured confounding.	This question should be answered after all studies have been considered, so a judgement can be made on how different domains could influence the estimate of the outcome, e.g. is there likely to be an impact of even a single uncontrolled domain on the outcome?	
BIAS IN PARTICIPANT SELECTION			
2.1 Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	NA/Y/PY/PN/N/NI This domain is concerned only with selection into the study based on participant characteristics observed after the start of intervention. Selection based on characteristics observed before the start of intervention can be addressed by controlling for imbalances between experimental intervention and		Were women included in the study based on any characteristics measured after their hyperglycaemic status was determined?
Controlling for post-intervention variables that are mediating variables estimates the common effects of intervention	Intervention is hyperglycaemia measured after the diagnosis was should not have been controlled for	affected by intervention is not appropriate. direct effect of intervention and may introduce and outcome introduces bias.	category, so any variables made (e.g. neonatal outcomes) Controlling for bias. Controlling for

– Screening for Gestational Diabetes

Question	Response options	Literature-recommended criteria	Criteria for GDM Rapid Review
2.2 If Y/PY to 2.1: Were the postintervention variables that influenced selection likely to be associated with intervention? and, 2.3 If Y/PY to 2.2: Were the postintervention variables that influenced selection likely to be outcome or a cause	baseline characteristics that are prognostic for the outcome (baseline confounding). NA/Y/PY/PN/N/NI influenced by the	If selection was also dependent on other inclusion outcome observed, e.g. macrosomia, LGA? criteria, were these associated with the intervention and if so, were these able to influence	Was the decision to include a woman in the analysis based on inclusion criteria that may have also been associated with glucose levels, e.g. BMI or ethnicity or with the
2.4 Do start of follow-up and start of intervention coincide for most participants?	NA/Y/PY/PN/N/NI	If participants are not followed from the start of the intervention then a period of follow up has been excluded, and individuals who experienced the outcome soon after intervention will be missing from analyses. This problem may occur when prevalent, rather than new (incident), users of the intervention are included in analyses.	Did all women have their hypoglycaemia diagnosed around the same time and were followed up until birth?
of the outcome? comparator groups in 2.5 If Y/PY to 2.2 and 2.3, or N/PN to	Questions combined due to high similarity. NA/Y/PY/PN/N/NI 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?	the outcome?	

Response options	Literature-recommended criteria	Criteria for GDM Rapid Review	
It is in principle possible to correct for selection biases, for example by using inverse probability weights to create a pseudo-population in which the selection bias has been removed, or by modelling the distributions of the missing participants or follow up times and outcome events and including them using missing data			
Risk of bias judgement	Low/ Moderate/ Serious/ Critical/ NI **Low**: (i) All participants who would have been eligible for the target trial were included in the study; and (ii) For each participant, start of follow up and start of intervention coincided. **Moderate**: (i) Selection into the study may have been related to intervention and outcome but appropriate methods to adjust for selection bias used or (ii) Start of follow up and start of intervention do not coincide for all participants; and either the proportion of participants for which this was the case was low or appropriate methods were used to account for this or it can be said with confidence the effect of intervention remains constant over time. **Serious**: (i) Selection into the study was		
methodology. However such methods are rarely used and the answer to this question will usually be “No”.			
	related (but not very strongly) to intervention and outcome, and could not be adjusted for in analyses; or (ii) Start of follow up and start of intervention do not coincide and a potentially important amount of follow-up time is missing from analyses and the rate ratio is not constant over time **Critical**: (i) Selection into the study was very strongly related to intervention and outcome; and could not be adjusted for in analyses or (ii) A substantial amount of follow-up time is likely to be missing from analyses and the rate ratio is not constant over time.		
BIAS IN THE CLASSIFIC ATION OF INTERVENTIONS			
3.1 Were intervention groups clearly defined?	NA/Y/PY/PN/N/NI	A pre-requisite for an appropriate comparison of interventions is that the interventions are well defined. Ambiguity in the definition may lead to bias in the classification of participants. For individuallevel interventions, criteria for considering individuals to have received each intervention should be clear and explicit, covering issues such as type, setting, dose, frequency, intensity and/or timing of intervention.	Were the criteria for a woman classified as hyperglycaemic/GDM clear?

– Screening for Gestational Diabetes

3.2 Was the information used to define NA/Y/PY/PN/N/NI In general, if i information about interventions intervention groups recorded at the start c received is available from sources that could not a

of the intervention? have been affected by subsequent outcomes, then t d i i o f n f o e f r i e n n

t t e i r a v l e m n i t s i c o l n a s s t s a i

t f u

s i i n s t u e n r l v i e k n e t l i y o

. n

C m o a l k l e e s c i t t i e o a n s o i f e t r h t e o i a n v f o o i r

d m s a u t c i h o m n i a s t c t l h a e s t s i i m f e i o c f a t t h

i e o

3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	If a retrospective study is evaluating lifestyle interventions (e.g. diet or diet + exercise, or different dietary modifications), it may be possible that patients may be misclassified due the more fluid nature of the intervention	Collection of the information at the time of the intervention may not be sufficient to avoid bias. The way in which the data are collected for the purposes of the NRSI should also avoid misclassification.

Was the test for glucose levels that determined classification into hyperglycaemia done before any other outcomes were collected, e.g. could any women have already been suspected to have higher blood glucose/GDM and is there therefore a risk that women were classified into hyperglycaemic because of outcome, rather than glucose levels?

n .

o f i n t e r v e n t i o n . C o i n t

4.3. Were important co-interventions NA/Y/PY/PN/N/NI Risk of bias will be higher if unplanned cobalanced across intervention groups?

interventions were implemented in a way that

w o u l d b i a s

Risk of bias judgement	Low/ Moderate/ Serious/ Critical/ NI	
Response options	**Literature-recommended criteria**	**Criteria for GDM Rapid Review**
	Low: Intervention status is well defined and the definition is based solely on information collected at the time of intervention. **Moderate**: Intervention status is well defined but some aspects of the	Was there any risk that women who were hyperglycaemic were treated differently than those normal? E.g. if the study was based on reclassifying women based on medical records, then it's unlikely they were treated differently. In prospective studies, those with hyperglycaemia might have been more "conscious" of their state. Consider if that is somewhat at a risk of bias?

Question	Response options	Literature-recommended criteria	Criteria for GDM Rapid Review
	assignments of intervention status were determined retrospectively. **Serious**: (i) Intervention status is not well defined; or (ii) Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. **Critical**: (Unusual) An extremely high amount of misclassification of intervention status, e.g. because of unusually strong recall biases.		

BIAS DUE TO DEVIATIONS FROM INTENDED INTERVENTIONS

– Screening fo

BIAS DUE TO MISSING DATA

Risk of bias judgement	Low/ Moderate/ Serious/ Critical/ NI Low: (i) Any deviations from intended intervention reflected usual practice; or (ii) Any deviations from usual practice were unlikely to impact on the outcome. Moderate: There were deviations from usual practice, but their impact on the outcome is expected to be slight. Serious: There were deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome. Critical: There were substantial deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome.		
5.1 Were outcome data available for all, or nearly all, participants?	NA/Y/PY/PN/N/NI	"Nearly all" should be interpreted as "enough to be confident of the findings", and a suitable prop ortion depends on the context. In some situations, availability of data from 95% (or possibly 90%) of the participants may be sufficient, providing th at events of interest are reasonably common in both intervention groups. One aspect of this is that review authors would ideally try and locate an analysis plan for the study.	Were at least 80% women who were considered for the study included in the final analysis?
5.2 Were participants excluded due to missing intervention status?	NA/Y/PY/PN/N/NI	Missing intervention status may be a problem. This requires that the intended study sample is clear, they did not have a measurement of incomplete data on aspects of intervention which it may not be in practice.	Were any women excluded because they did not have a measurement of glucose levels? In e.g. chart review, participants with incomplete data on aspects of intervention received might be removed from analysis.
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	NA/Y/PY/PN/N/NI	This question relates particularly to participants excluded from the analysis because of missing information on confounders that were controlled for in the analysis.	Were any women excluded from the analysis because they had missing data other than on variable that were adjusted for/needed for analysis (e.g. age, BMI)?
5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?	NA/Y/PY/PN/N/NI	This aims to elicit whether either (i) differential proportion of missing observations or (ii) differences in reasons for missing observations could substantially impact on our ability to answer the question being addressed. "Similar" includes some minor degree of discrepancy across intervention groups as expected by chance.	Was there a difference between how many hyperglycaemic/GDM/normal women were excluded or were they excluded for different reasons.

Response options	Literature-recommended criteria	Criteria for GDM Rapid Review	
5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	NA/Y/PY/PN/N/NI	Evidence for robustness may come from how missing data were handled in the analysis and whether sensitivity analyses were performed by the investigators, or occasionally from additional analyses performed by the systematic reviewers. It is important to assess whether assumptions employed in analyses are clear and plausible. Both content knowledge and statistical expertise will often be required for this. For instance, use of a statistical method such as multiple imputation does not guarantee an appropriate answer. Review authors should seek naïve (completecase) analyses for comparison, and clear differences between complete-case and multiple imputationbased findings should lead to careful assessment of the validity of the methods used.	Is it likely that women who were excluded from the analysis were different from the rest, i.e. is it likely that they had less or more severe outcomes? Would their inclusion likely influence the results?
Risk of bias judgement	Low/ Moderate/ Serious/ Critical/ NI **Low**: (i) Data were reasonably complete; or (ii) Proportions of and reasons for missing participants were similar across intervention groups; or (iii) The analysis addressed missing data and is likely to have removed any risk of bias. **Moderate**: (i) Proportions of and reasons for missing participants differ slightly across intervention groups; and (ii) The analysis is unlikely to have removed the risk of bias arising from the missing data. **Serious**: (i) Proportions of missing participants or reasons for missingness differ substantially across interventions		

Question	Response options	Literature-recommended criteria	Criteria for GDM Rapid Review
	and (ii) The analysis is unlikely to have removed the risk of bias arising from the missing data or missing data were addressed inappropriately in the analysis or the nature of the missing data means that the risk of bias cannot be removed through appropriate analysis. **Critical**: (unusual) There were critical difference between interventions in participants with missing data and missing data were not (or could not) be addressed through appropriate analysis		
BIAS IN MEASUREMENT OF OUTCOMES			

6.2 Were outcome assessors aware of the intervention received by study participants?	NA/Y/PY/PN/N/NI I.e. was the person collecting the outcome data "blind" to the intervention? Question to be answered for each relevant reported outcome separately	If outcome assessors were blinded to intervention status, the answer to this question would be 'No'. In other situations, outcome assessors may be unaware of the interventions being received by participants despite there being no active blinding by the study investigators; the answer this question would then also be 'No'. In studies where participants report their outcomes themselves, for example in a questionnaire, the outcome assessor is the study participant. In an observational study, the answer to this question will usually be 'Yes' when the participants report their outcomes themselves.

been influenced
by 6.1 Could the outcome measures have NA/Y/PY/PN/N/NI knowledge of
the
I.e. would it matter if the study wasn't
intervention received? "blinded"? Question to be answered for each
relevant reported outcome separately. Some outcome measures involve
negligible Were the obstetricians/midwives assessor judgment, e.g. all-
cause mortality or non aware of the women's status as repeatable automated
laboratory assessments. normal/hyperglycemic/GDM? Risk of bias due to
measurement of these outcomes would be expected to be low.
For retrospective chart analyses/database studies, where women e.g. re-classified
using different criteria, this should be a

outcome assessment NA/Y/PY/PN/N/NI
comparable across
intervention groups?
"no". For prospective studies,
consider if the glucose levels
would have been considered by
the delivery personnel or people
who measure the outcomes

6.3 Were the
methods of

Response options	Literature-recommended criteria	Criteria for GDM Rapid Review	
This is likely to be similar, but might actually differ if some patients were treated much later than others. It is mostly expected to be different for studies where a historical control is used or in beforeandafter studies			
Comparable assessment methods (i.e. data Were outcomes assessed in the collection) would involve the same outcome same manner for all women? E.g. detection methods and thresholds, same time point, was large for gestational age same definition, and same measurements. determined in the same was for hyperglycaemic and normal women?	6.4 Were any systematic errors in measurement of the outcome related to intervention received?	NA/Y/PY/PN/N/NI Unless the outcomes were measured with different methods between the intervention groups, this is highly unlikely.	This question refers to differential misclassification of outcomes. Systematic errors in measuring the outcome, if present, could cause bias if they are related to intervention or to a confounder of the intervention-outcome relationship. This will usually be due either to outcome assessors being aware of the intervention received or to non-comparability of outcome assessment methods, but there are examples of differential misclassification arising despite these controls being in place.	Could outcome measurement be biased based on the presence/absence of hyperglycaemia, e.g. were hypoglycaemic women checked for the outcome more often than “normal glucose” women?

e

w – Screening for Gestational Diabetes

Risk of bias judgement	Low/ Moderate/ Serious/ Critical/ NI **Low**: The methods of outcome assessment were comparable across intervention groups and either the outcome measure was unlikely to be influenced by lack of blinding or outcome assessors were unaware of intervention received by study participants and error in outcome measurement is unrelated to intervention status. **Moderate**: (i) The methods of outcome assessment were comparable across intervention groups; and (ii) The outcome measure is only minimally influenced by knowledge of the intervention received by study participants; and (iii) Any error in measuring the outcome is only minimally related to intervention status. **Serious**: (i) The methods of outcome assessment were not comparable across intervention groups; or (ii) The outcome measure was subjective and the outcome assessed by assessors aware of the intervention received by study participants; or (iii) Error in measuring the outcome was related to intervention status. **Critical**: The methods of outcome assessment were so different that they cannot reasonably be compared across intervention groups.

BIAS IN SELECTION OF THE REPORTED RESULT			
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome measurements within the outcome domain?	NA/Y/PY/PN/N/NI E.g. if the number of people with a specific threshold of the outcome is reported, the threshold can be changed. This should be easy to spot if the threshold is different than what is usually used in other studies or guidelines. If multiple glucose tests are administered to measure postpartum glucose level or	For a specified outcome domain, it is possible to generate multiple effect estimates for different measurements. If multiple measurements were made, but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Is there suspicion that the outcome was measured in different ways and then preferentially reported? E.g. the authors measured birthweight as a continuous variable but then only reported number of babies below/above a certain threshold?

Question	Response options	Literature-recommended criteria	Criteria for GDM Rapid Review
	diabetes, and data are only reported for one particular test result, this could increase risk of bias.		
.2 ... multiple analyses of the interventionoutcome relationship?	NA/Y/PY/PN/N/NI	Because of the limitations of using data from nonrandomized studies for analyses of effectiveness (need to control confounding, substantial missing data, etc), analysts may implement different analytic methods to address these limitations. Examples include unadjusted and adjusted models; use of final value vs change from baseline vs analysis of covariance; different transformations of variables; a continuously scaled outcome converted to categorical data with different cutpoints; different sets of covariates used for adjustment; and different analytic strategies for dealing with missing data. Application of such methods generates multiple estimates of the effect of the intervention versus the comparator on the outcome. If the analyst does not pre-specify the methods to be applied, and multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	If adjustments for confounding variables/regressions were done, were the analyses prespecified or is there reason to believe that these were not prespecified but selected based on outcome of the analyses?
7.3 ... different subgroups?	NA/Y/PY/PN/N/NI	Particularly with large cohorts often available from analyses based on routine data sources, it is possible to generate multiple effect estimates for different subgroups or simply to omit varying proportions of the original cohort. If multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Were women subgroups only?

Risk of bias judgement	Low/ Moderate/ Serious/ Critical/ NI **Low**: There is clear evidence (usually through examination of a pre-registered protocol or statistical analysis plan) that all reported results correspond to all intended outcomes, analyses and sub-cohorts. **Moderate**: (i) The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and (ii) There is no indication of selection of the reported analysis from among multiple analyses; and (iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results. **Serious**: (i) Outcomes are defined in different ways in the methods and results sections, or in different publications of the study; or (ii) There is a high risk of selective reporting from among multiple analyses; or (iii) The cohort or subgroup is selected from a larger study for analysis and appears to be reported on the basis of the results. **Critical**: (i) There is evidence or strong suspicion of selective reporting of results; and (ii) The unreported results are likely to be substantially different from the reported results.
OVERALL BIAS	Low/ Moderate/ Serious/ Critical/ NI Low: The study is judged to be at low risk of bias for all domains. Moderate: The study is judged to be at low or moderate risk of bias for all domains. Serious: The study is judged to be at serious risk of bias in at least one domain, but not at critical risk of bias in any domain. Critical: The study is judged to be at critical risk of bias in at least one domain.

Table 98. Guidance on the use of AMSTAR 2

Screening for Gestational Diabetes

Question	Response options	Literature-recommended criteria	Criteria for GDM Rapid Review

Did the review authors use a comprehensive literature search strategy? (Yes/Partial Yes/No)	To score Yes, appraisers should be satisfied that all relevant aspects of the search have been addressed by review authors
Did the review authors perform study selection in duplicate? (Yes/No)	If one reviewer carried out selection of all studies with a second reviewer checking agreement on a sample of studies, a Kappa score indicating 'strong' agreement (≥0.80) should have been achieved
Did the review authors perform data extraction in duplicate? (Yes/No)	If one reviewer carried out extraction of all studies with a second reviewer checking agreement on a sample of studies, a Kappa score indicating 'strong' agreement (≥0.80) should have been achieved
Did the review authors provide a list of excluded studies and justify the exclusions? (Yes/Partial Yes/No)	Exclusion should not be based on RoB, which is dealt with separately and later in the review process

Question	**Literature-Recommended Criteria**
	Please refer to the PDF checklist for the full criteria that must be fulfilled to receive a Yes or Partial Yes answer. Additional pointers for each question are described in the full guidance but key details have been pulled out below
Did the research questions and inclusion criteria for the review include the components of PICO? (Yes/No)	To score Yes, appraisers should be confident that the 4 elements of PICO are described somewhere in the report
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol? (Yes/Partial Yes/No)	The research questions and study methods should have been planned ahead of conducting the review (this should be reported at minimum to score a Partial Yes) To score Yes, authors should demonstrate that they worked with a written protocol with independent verification (e.g. in the form of registration, an open publication journal or a date submission to a research office or research ethics board). Appraisers should compare the published review report with the registered protocol (if available); if there are deviations from the protocol, the appraisers should determine whether these are reported and justified by the review authors. Obvious unexplained discrepancies should result in downgrading the rating
Did the review authors explain their selection of the study designs for inclusion in the review? (Yes/No)	The justification for selection of study designs may have to be inferred from careful reading of the complete study report The general rule is that authors first asked whether a review restricted to RCTs would have given an incomplete summary. If the answer to this is yes, the inclusion of non-randomised studies is justified Restriction to only non-randomised studies is justified when RCTs will not provide the necessary outcome data, or if a review of RCTs has already been completed and the aim is to complement this Inclusion of both RCTs and non-randomised studies may be justified to get a complete picture; in this situation it is recommended that the two study types are assessed and combined independently

Question	**Literature-Recommended Criteria**
Did the review authors describe the included studies in adequate detail? (Yes/Partial Yes/No)	The detail should be sufficient for an appraiser, or user, to make judgements about the extent to which the studies were appropriately chosen (in relation to the PICO structure)
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review? (Yes/Partial Yes/No)	When the review is confined to RCTs, it is recommended that the Cochrane Handbook is consulted to determine whether review authors made an adequate assessment of RoB in individual RCTs Review authors should have used a systematic approach to RoB assessment, preferably with a properly-developed rating instrument (if they have used a non-standard instrument the appraiser should be satisfied that it was capable of detecting serious methodological flaws) In assessing how RoB has been assessed by review authors it is recommended that appraisers should seek methods and content expert advice (if that is not included in the team), along with guidance on what adjustment techniques for confounding would be appropriate The domains of bias selected from the ROBINS-I instrument as being the most relevant to SLRs that include non-randomised studies of interventions include: confounding, sample selection bias, bias in measurement of exposures and outcomes, selective reporting of outcomes and analyses

Did the review authors report on the sources of funding for the studies included in the review? (Yes/No)	No additional guidance

If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results? (Yes/No/No metaanalysis conducted) Review authors should have stated explicitly in the review protocol the basis of their decision to perform a meta-analysis e.g. desire to obtain a	single pooled effect and the extent to which studies are able to be combined Authors should have explained decisions to use fixed or random effects models (for RCTs) and the methods they intended to use to investigate heterogeneity Pooled estimates should be reported separately for different study types (i.e. not combining RCTs and non-randomised studies of interventions) For non-randomised studies of interventions, authors should pool the confounder-adjusted estimates of effect rather than raw data (there should be a clear justification if they do the latter). N.B. different studies are likely to report treatment effects that have been adjusted for different sets of covariates – another source of potential heterogeneity

Screening for Gestational Diabetes

Question	Response options	Literature-recommended criteria	Criteria for GDM Rapid Review
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the metaanalysis or other evidence synthesis? (Yes/No/No meta-analysis conducted)	This is particularly important where the review includes RCTs of variable quality. The impact of this should be assessed by regression analysis or by estimating pooled effect sizes with only studies at low RoB For non-randomised studies of interventions, they should estimate pooled effect sizes of low/moderate RoB studies If meta-analyses were not performed, the authors should still comment on the likely impact of RoB on individual study results (see next item)		
Did the review authors account for RoB in individual studies when interpreting/discussing	This discussion should not be limited to the impact of RoB on pooled estimates, but should also consider whether it may account for differences between the results of individual studies **the results of the review? (Yes/No)** The authors should make an explicit consideration of RoB if they make any recommendations that are likely to impact clinical care or policy		
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? (Yes/No)	Both the PICO elements and domains of bias (listed in item 9) should be considered as potential sources of heterogeneity in the results Review authors should explore these and discuss the impact of heterogeneity on the results, conclusions and any recommendations		
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review? (Yes/No/No meta-analysis conducted)	This can be a difficult issue to resolve. The key issues are whether review authors have tried to identify publication bias through additional literature searches, shown an awareness of the likely impact of publication bias in their interpretation and discussion or results, and performed a sensitivity analysis to determine how many missing 'null' studies (i.e. those not published because of an insignificant result) would be needed to invalidate the results of the SLR Typically, statistical tests/graphic displays are used and if they are positive it indicates the presence of publication bias, however negative tests do not guarantee its absence as the tests are insensitive Context and setting should also be considered (e.g. a series of industry-sponsored studies may be more likely to be affected by publication bias than similar studies independent of industry)		
Did the review authors report any potential sources of conflict of interest, including any funding they received for the review? (Yes/No)	No additional guidance		

Table 99. Guidance on the use of QUADAS-2

Question	Literature-Recommended Criteria	Guideline Criteria for Gestational Diabetes Studies

PARTICIPANT SELECTION

Was a consecutive or random sample of pregnancies enrolled? A study should ideally enrol all consecutive, or a random sample of eligible patients – otherwise there is potential for bias. Studies that make inappropriate exclusions, e.g. excluding "difficult to diagnose" may result in overoptimistic estimates of diagnostic accuracy

Yes if all pregnancies (or a random sample of patients) within the study period were included

No if patients were selected in a different way, e.g. by referral or convenience sample

		Unclear if all screened pregnancies are enrolled but it is not specified if the screening test is routinely administered at the study site
Was a case-control design avoided?	Studies enrolling patients with known disease and a control group without the condition may exaggerate diagnostic accuracy	**Yes** if the study was a prospective or retrospective cohort study **No** if cases (gestational diabetes) were matched to controls
Did the study avoid inappropriate exclusions?	Exclusion of patients with "red flags" for the target condition, who may be easier to diagnose, may lead to underestimation of diagnostic accuracy	**Yes** if all pregnancies were included, or if exclusions were appropriate and unlikely to lead to bias **No** if any group within the screening population was systematically excluded
Could the selection of pregnancies have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the included pregnancies do not match the review question?	There may be concerns regarding applicability if patients included in the study differ, compared to those targeted by the review question, in terms of severity of the target condition, demographic features, presence of differential diagnosis or co-morbidity, setting of the study and previous testing protocols	**Low** if patients overall are low-risk pregnancies representative of the screening population (i.e. similar to the pregnant population in the UK) **High** if patients overall are not representative of the screening population, such as pregnancies with at least one moderate risk factor as specified in UK guidelines or demographically dissimilar to the UK population
INDEX TESTS		
Were the index test results Interpretation **interpreted without knowledge of** of the **the reference standard?**	This item is similar to "blinding" in intervention studies. Interpretation of index test results may be influenced by knowledge of reference standard	**Yes** if screening results were interpreted before the diagnosis was confirmed **No** if screening results were only examined after the diagnosis was confirmed
If a threshold was used, was it pre-specified?	Selecting the test threshold to optimise sensitivity and/or specificity may lead to overoptimistic estimates of test performance, which is likely to be poorer in an independent sample of patients in whom the same threshold is used	**Yes** if the criteria used to diagnose gestational diabetes were explicitly stated, well-defined, and specified before the study **No** if criteria were not stated, were insufficiently well-defined, or were specified retrospectively
Could the conduct or interpretation of the index test introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain. Consider whether the staff conducting the index test could have had **have** foreknowledge of who was at risk by presence of major factors.
Is there concern that the index test, its conduct, or interpretation differ from the review question?	Variations in test technology, execution, or interpretation may affect estimates of its diagnostic accuracy. If index tests methods vary from those specified in the review question there may be concerns regarding applicability	**Low** if the screening test is similar to tests or screening tests administered as part of UK clinical practice **High** if any aspect of the index test, including its conduct or interpretation, is substantially different from clinical practice in a UK setting (as outlined in the NG3 NICE guidance)

REFERENCE STANDARD		
Is the reference standard likely to correctly classify the test condition?	Estimates of test accuracy are based on the assumption that the reference standard is 100% sensitive and specific. Disagreements between the reference standard and index test are assumed to result from incorrect classification by the index test	**Yes** if gestational diabetes was confirmed consistently at ≥24 completed weeks of gestation based an accepted definition (by any relevant guideline) **No** if diagnosis was performed inconsistently, or if the methods used are likely to be unreliable
Were the reference standard interpreted without knowledge of the results of the	Potential for bias is related to the potential influence of prior knowledge on the interpretation of the reference standard	**Yes** if the final diagnosis of gestational diabetes was made by an **results** investigator blinded to the index test results **No** if the screening results were known by the investigator making the **index test?** final diagnosis **Unclear** if it is not clear whether the investigator was aware of the test result when making the final diagnosis
Could the reference standard, its conduct, or its interpretation have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the target condition as defined by the reference standard does not match the review question? The reference standard may be free of bias but the target **Was there an appropriate interval between the index test(s) and**	condition that it defines may differ from the target condition specified in the review question. For example, when defining urinary tract infection, the reference standard is generally based on specimen culture but the threshold above which a result is considered positive may vary **Low** if the definition of gestational diabetes used was the standard UK definition or similar: are collected on the same patients at the same time. If there is a delay or if treatment is started between index test and reference	FPG of ≥5.6mmol/L, or a 2-hour 75 g OGTT plasma glucose level of ≥7.8mmol/L OR **High** if the reference standard defined gestational diabetes in any other way
Did all participants receive a reference standard? **Did participants receive the same reference standard?**	Verification bias occurs when not all of the study group receive confirmation of the diagnosis by the same reference standard. If the results of the index test influence the decision on whether to perform the reference standard or which reference standard is used, estimated diagnostic accuracy may be biased	**Yes** if all screened patients had confirmation of their diagnosis, and all were diagnosed in the same manner (using the same reference standard by similarly trained staff) **No** if patients received different reference standards **Unclear** if there was a high variability in staff diagnosing and recording pre-eclampsia, or the staff may not have received the same training
Were all pregnancies included in the analysis?	All patients who were recruited into the study should be included in the analysis. There is a potential for bias if the number of patients enrolled differs from the number of patients included in the 2x2 table of results, for example because patients lost to follow-up differ systematically from those who remain	**Yes** if all screened women were included in the final analysis **No** if any screened women were not included in the final analysis
the reference standard? **PARTICIPANT FLOW** Ideally results of the index test and reference standard	standard, misclassification may occur due to recovery or deterioration of the condition. The length of interval leading to a high risk of bias will vary between conditions. A delay of a few days may clinical not be a problem for chronic conditions, while for acute infectious diseases a short delay may be important	**Yes** if the reference standard was given before any change in care occurred (e.g. treatment) **No** if some women were treated to avoid developing gestational diabetes

Could the participant flow have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	**No** if women who underwent the index test were all equally likely to develop and be diagnosed with gestational diabetes in the same manner **Yes** if some women could have been prevented from developing gestational diabetes (e.g. by labour induction) or if women received different reference standards or a significant proportion were removed from the analysis

Appendix 5 – Appraisal for quality and risk of bias

Table 100. Summary of the AMSTAR 2 assessment of the Farrar 2016 SLR investigating outcomes in GDM

Question	Farrar 2016[4]
Did the research questions and inclusion criteria for the review include the components of PICO? (Yes/No)	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol? (Yes/Partial Yes/No)	Yes
Did the review authors explain their selection of the study designs for inclusion in the review? (Yes/No)	No
Did the review authors use a comprehensive literature search strategy? (Yes/Partial Yes/No)	Yes
Did the review authors perform study selection in duplicate? (Yes/No)	Yes
Did the review authors perform data extraction in duplicate? (Yes/No)	Yes
Did the review authors provide a list of excluded studies and justify the exclusions? (Yes/Partial Yes/No)	Yes
Did the review authors describe the included studies in adequate detail? (Yes/Partial Yes/No)	Yes
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review? (Yes/Partial Yes/No)	Yes
Did the review authors report on the sources of funding for the studies included in the review? (Yes/No)	Yes
If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results? (Yes/No/No meta-analysis conducted)	Yes
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis? (Yes/No/No meta-analysis conducted)	Yes
Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review? (Yes/No)	Yes

Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? (Yes/No)	Yes
	Page
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review? (Yes/No/No meta-analysis conducted)	No
Did the review authors report any potential sources of conflict of interest, including any funding they received for the review? (Yes/No)	Yes

Table 101. Summary of the AMSTAR 2 assessment the Farrar 2016 SLR investigating screening for GDM

Question	Farrar 2016, Chapter 5.2
Did the research questions and inclusion criteria for the review include the components of PICO? (Yes/No)	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol? (Yes/Partial Yes/No)	Partial yes
Did the review authors explain their selection of the study designs for inclusion in the review? (Yes/No)	No
Did the review authors use a comprehensive literature search strategy? (Yes/Partial Yes/No)	Partial yes
Did the review authors perform study selection in duplicate? (Yes/No)	Yes
Did the review authors perform data extraction in duplicate? (Yes/No)	Yes
Did the review authors provide a list of excluded studies and justify the exclusions? (Yes/Partial Yes/No)	Yes
Did the review authors describe the included studies in adequate detail? (Yes/Partial Yes/No)	Yes
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review? (Yes/Partial Yes/No)	Partial yes
Did the review authors report on the sources of funding for the studies included in the review? (Yes/No)	No
If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results? (Yes/No/No meta-analysis conducted)	No meta-analysis conducted
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis? (Yes/No/No meta-analysis conducted)	No meta-analysis conducted
Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review? (Yes/No)	No

Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? (Yes/No)	No
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review? (Yes/No/No meta-analysis conducted)	No meta-analysis conducted
Did the review authors report any potential sources of conflict of interest, including any funding they received for the review? (Yes/No)	Yes

Table 102. Summary of AMSTAR 2 assessments for SLRs evaluating treatment of pregnant women with GDM

Question	6) Farrar 2016[4] Chapter	Brown 2017L lifestyle interventions)[91] (	) Brown 2017I Insulin SLR[92]	Brown 2017A anti-diabetic (medication)[93]
Did the research questions and inclusion criteria for the review include the components of PICO? (Yes/No)	Yes	Yes	Yes	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol? (Yes/Partial Yes/No)	Yes	Yes	Yes	Yes
Did the review authors explain their selection of the study designs for inclusion in the review? (Yes/No)	No	No	No	No
Did the review authors use a comprehensive literature search strategy? (Yes/Partial Yes/No)	Yes	Yes	Yes	Yes
Did the review authors perform study selection in duplicate? (Yes/No)	Yes	Yes	Yes	Yes
Did the review authors perform data extraction in duplicate? (Yes/No)	Yes	Yes	Yes	Yes
Did the review authors provide a list of excluded studies and justify the exclusions? (Yes/Partial Yes/No)	Yes	Yes	Yes	Yes
Did the review authors describe the included studies in adequate detail? (Yes/Partial Yes/No)	Yes	Yes	Yes	Yes
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review? (Yes/Partial Yes/No)	Yes	Yes	Yes	Yes
Did the review authors report on the sources of funding for the studies included in the review? (Yes/No)	No	Yes	Yes	Yes
If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results? (Yes/No/No meta-analysis conducted)	Yes	Yes	Yes	Yes
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis? (Yes/No/No meta-analysis conducted)	Yes	Yes	Yes	Yes
Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review? (Yes/No)	Yes	Yes	Yes	Yes
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? (Yes/No)	Yes	Yes	Yes	Yes

Question	6) Farrar 2016 [4] Chapter	Brown 2017L lifestyle interventions)[91] (	) Brown 2017I Insulin SLR [92]	Brown 2017A anti-diabetic (medication)[93]
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review? (Yes/No/No meta-analysis conducted)	(Yes	No	Yes (	No
Did the review authors report any potential sources of conflict of interest, including any funding they received for the review? (Yes/No)	Yes	Yes	Yes	Yes

Table 103: Cochrane ROBINS-I

Question	Beksac 2018	Berggren 2011	Berggren 2012 (MFMU network)	Berntorp 2015 (Mamma study)	Biri 2009	Cheng 2009
BIAS DUE TO CONFOUNDING						
1.1 Is there potential for confounding of the effect of intervention in this study?	PY	PY	Y	PY	PY	PY
1.4 Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	Y	Y	Y	Y	N	Y
1.5 If Y/PY to 1.4 Were confounding domains that were controlled for measured	Y	PY	Y	Y	NA	PY validly and reliably by the variables available in this study?
1.6 Did the authors control for any post-intervention variables that could have	N	N	been affected by the	N	N	Y
intervention?						
Risk of bias judgement	Moderate	Moderate	Moderate	Moderate	Serious	Serious
BIAS IN PARTICIPANT SELECTION						

observed after the start of intervention?	N	N characteristics		N	Y	Y
2.1 Was selection of participants into the study (or into the analysis) based on		participant	PN			
2.2 If Y/PY to 2.1: Were the post-intervention variables that influenced selection	NA	NA	NA	NA	PN	PN

Question	Beksac 2018	Berggren 2011	Berggren 2012 (MFMU network)	Berntorp 2015 (Mamma study)	Biri 2009	Cheng 2009
likely to be associated with intervention? and, 2.3 If Y/PY to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?						
2.4 Do start of follow-up and start of intervention coincide for most participants?	Y	Y	Y	Y	Y	Y
2.5 If Y/PY to 2.2 and 2.3, or N/PN to 2.4: Were adjustment techniques used the presence of selection biases?	NA	NA	NA that are likely to correct for	NA	NA	NA
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS IN THE CLASSIFICATION OF INTERVENTIONS						
3.1 Were intervention groups clearly defined?	Y	Y	Y	Y	Y	Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	Y	Y	Y	Y	Y	Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	N	N	N	N	N	N
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS DUE TO DEVIATIONS FROM INTENDED INTERVENTIONS						
4.3. Were important co-interventions balanced across intervention groups?	Y	N	Y	PY	N	N
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS DUE TO MISSING DATA						
5.1 Were outcome data available for all, or nearly all, participants?	PY	Y	Y	Y	PY	Y
5.2 Were participants excluded due to missing data on intervention status?	PN	N	N	PN	N	N

5.3 Were participants excluded due to missing data on other variables needed for the analysis?	PN	N	N	Y	N	N
5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?	NA	NA	NA	NI	NA	NA
5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	NA	NA	NA	PY	NA	NA
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS IN MEASUREMENT OF OUTCOMES						

Question	Beksac 2018	Berggren 2011	Berggren 2012 (MFMU network)	Berntorp 2015 (Mamma study)	Biri 2009	Cheng 2009
6.1 Could the outcome measures have been influenced by knowledge of the intervention received?	PN	PY		PN	PN	PN
6.2 Were outcome assessors aware of the intervention received by study participants?	PY	PY	PY		PY	PY
6.3 Were the methods of outcome assessment comparable across intervention groups?	PY	PY	PY	PY	PY	PY
6.4 Were any systematic errors in measurement of the outcome related to intervention received?	PY	PN	PY	PY	PN	PN
Risk of bias judgement	Moderate	Low	Moderate	Moderate	Low	Low
BIAS IN SELECTION OF THE REPORTED RESULT						
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome measurements within the outcome domain?	N	PN	N	N	PN	PN
.2 ... multiple analyses of the intervention-outcome relationship?	PN	PN	PY	PY	PN	PN
7.3 ... different subgroups?	N	PN	N	N	PN	PN

Risk of bias judgement	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate
OVERALL RISK OF BIAS	**Moderate**	**Moderate**	**Moderate**	**Moderate**	**Serious**	**Moderate**

Table 104: Cochrane ROBINS-I

Question	Corrado 2009	Delibas 2018	Davis 2018 (1290)	Donovan 2017	Ezell 2015	Jiang 2017
BIAS DUE TO CONFOUNDING						
1.1 Is there potential for confounding of the effect of intervention in this study?	PY	PY	PY	PY	Y	PY
1.4 Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	N	N	Y	Y	Y	Y
1.5 If Y/PY to 1.4 Were confounding domains that were controlled for measured	NA	NA	Y	PY	PY	PY validly and reliably by the variables available in this study?
1.6 Did the authors control for any post-intervention variables that could have been affected by the intervention?	N	N	N	N	N	N

Question	Corrado 2009	Delibas 2018	Davis 2018 (1290)	Donovan 2017	Ezell 2015	Jiang 2017
Risk of bias judgement	Serious	Serious	Moderate	Moderate	Serious	Moderate
BI AS IN PARTICIPANT SELECTION						
2.1 Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	Y	N	N	N	N	N
2.2 If Y/PY to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention? and, 2.3 If Y/PY to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?	PN	NA	NA	NA	NA	NA
2.4 Do start of follow-up and start of intervention coincide for most participants?	Y	Y	Y	Y	PY	Y
2.5 If Y/PY to 2.2 and 2.3, or N/PN to 2.4: Were adjustment techniques used for the presence of selection biases?	NA	NA	NA that are likely to correct	NA	NA	NA

Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS IN THE CLASSIFICATION OF INTERVENTIONS						
3.1 Were intervention groups clearly defined?	Y	Y	Y	Y	PY	Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	Y	Y	Y	Y	Y	Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	N	N	N	N	N	N
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS DUE TO DEVIATIONS FROM INTENDED INTERVENTIONS						
4.3. Were important co-interventions balanced across intervention groups?	Y	PY	Y	PY	Y	Y
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS DUE TO MISSING DATA						
5.1 Were outcome data available for all, or nearly all, participants?	PY	Y	PY	PY	Y	PY
5.2 Were participants excluded due to missing data on intervention status?	N	N	Y	N	Y	PN
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	N	N	PN	Y	Y	PN
5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?	NA	NA	NI	N	NI	NA

Question	Corrado 2009	Delibas 2018	Davis 2018 (1290)	Donovan 2017	Ezell 2015	Jiang 2017
5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	NA	NA	NA	PY	PY	NA
Risk of bias judgement	Low	Low	Moderate	Low	Low	Low
BIAS IN MEASUREMENT OF OUTCOMES						

	PN	PN	PN	PY	PN	
6.1 Could the outcome measures have been influenced by knowledge of the PN intervention received?						
			PY	PY	PY	
6.2 Were outcome assessors aware of the intervention received by study PY participants?						
6.3 Were the methods of outcome assessment comparable across intervention PY groups?	PY	PY	PY	PY	PY	
6.4 Were any systematic errors in measurement of the outcome related to PN intervention received? PY	PY	PY	PN	PY		
Risk of bias judgement	Low	Moderate	Moderate	Moderate	Low	Moderate
BIAS IN SELECTION OF THE REPORTED RESULT						
Is the reported effect estimate likely to be selected, on the basis of the results, PN from...	N	N	N	N	N	
7.1. ... multiple outcome measurements within the outcome domain?						
.2 ... multiple analyses of the intervention-outcome relationship? PN	PY	PY	PN	PN	NI	
7.3 ... different subgroups? PN	N	N	N	N	N	
Risk of bias judgement	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate
OVERALL RISK OF BIAS	**Serious**	**Serious**	**Moderate**	**Moderate**	**Serious**	**Moderate**

Table 105: Cochrane ROBINS-I

Question	Jiang 2017	Lopez 2019	Meek 2015	Miyakoshi 2010	Noor 2019	Verd 2016
BIAS DUE TO CONFOUNDING						
1.1 Is there potential for confounding of the effect of intervention in this study? PY	PY	PY	PY	PY	PY	
1.4 Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	Y	Y	PY	Y	N	N

Question	Jiang 2017	Lopez 2019	Meek 2015	Miyakoshi 2010	Noor 2019	Verd 2016
1.5 If Y/PY to 1.4 Were confounding domains that were controlled for measured	PY	Y	PY	PY	NA	NA validly and reliably by the variables available in this study?
1.6 Did the authors control for any post-intervention variables that could have intervention?	N	N been affected by the		Y	N	N
Risk of bias judgement	Moderate	Moderate	Serious	Moderate	Moderate	Serious
BIAS IN PARTICIPANT SELECTION						
2.1 Was selection of participants into the study (or into the analysis) based on after the start of intervention?	N	N participant characteristics observed		N	N	Y	Y
2.2 If Y/PY to 2.1: Were the post-intervention variables that influenced selection to be associated with intervention?		NA likely	NA	NA	NA	PY	PY
and, 2.3 If Y/PY to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?						PN	PY
2.4 Do start of follow-up and start of intervention coincide for most participants?	Y		Y	Y	Y	NI	Y
2.5 If Y/PY to 2.2 and 2.3, or N/PN to 2.4: Were adjustment techniques used to correct for the presence of selection biases?	NA	that are likely	NA	NA	NA	N	N
Risk of bias judgement	Low	Low	Low	Low	Serious	Serious
BIAS IN THE CLASSIFICATIO N OF INTERVENTI ONS						
3.1 Were intervention groups clearly defined?	Y	Y	Y	Y	Y	Y
3.2 Was the information used to define intervention groups recorded at the start	Y	of the intervention?	Y	Y	Y	Y	Y
3.3 Could classification of intervention status have been affected by knowledge outcome or risk of the outcome?	N	of the	N	N	N	PN	N
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS DUE TO DEVIATIONS FROM INTENDED INTERVENTIONS						
4.3. Were important co-interventions balanced across intervention groups?	Y	Y	Y	Y	Y	Y
Risk of bias judgement	Low	Low	Low	Low	Low	Low

BIAS DUE TO MISSING DATA						
5.1 Were outcome data available for all, or nearly all, participants?	PY	PY	Y	PY	PY	Y
5.2 Were participants excluded due to missing data on intervention status?	PN	N	N	N	Y	N
Question	**Jiang 2017**	**Lopez 2019**	**Meek 2015**	**Miyakoshi 2010**	**Noor 2019**	**Verd 2016**
5.3 Were participants excluded due to missing data on other variables needed	PN	PN	Y for the analysis?	PN	Y	N
5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are the proportion of participants and similar across interventions?	NA	NA	NI reasons for missing data	NA	NI	NA
5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were missing data?	NA	NA	PN robust to the presence of	NA	Y	NA
Risk of bias judgement	Low	Low	Low	Low	Low	Low
BIAS IN MEASUREMEN OUTCOMES						
6.1 Could the outcome measures have been influenced by knowledge of the received?	PN	PN	PN intervention	PN	PY	PN
6.2 Were outcome assessors aware of the intervention received by study	PY	PY	PY participants?	PY	PY	PY
6.3 Were the methods of outcome assessment comparable across intervention	PY	PY	PY groups?	PY	PY	PY
6.4 Were any systematic errors in measurement of the outcome related to	PY	PY	PN intervention	PY	PN	PN
Risk of bias judgement	Moderate	Moderate	Low	Low	Low	Low
BIAS IN SELECTION OF THE REPORTED RE ULT						
Is the reported effect estimate likely to be selected, on the basis of the results, 7.1. ... multiple outcome measurements within the outcome domain?	N	N	PN from...	N	N	PN
.2 ... multiple analyses of the intervention-outcome relationship?	NI	NI	PN	NI	NA	PN
7.3 ... different subgroups?	N	N	PN	N	N	PN

Risk of bias judgement	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate
OVERALL RISK OF BIAS	**Moderate**	**Moderate**	**Moderate**	**Moderate**	**Serious**	**Serious**

Table 106: Cochrane RCTs

Question		Rowan 2018 (MiG)		Hernandez 2016		Senat 2018 (INDAO)	
RANDOMISATION PROCESS	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**	
1.1 Was the allocation sequence random?	PY	Randomisation carried out online – few details as this is a publication	NI	Not specified	secondary	Y	Computerrandomisation sequence -generated
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	See above	NI	Not specified		Y	Concealed to clinicians and participants
1.3 Did the baseline differences between intervention groups suggest a problem with the randomization process?	NI	Baseline characteristics not reported for full randomised cohort by arm	PN	Some differences, but likely to have arisen from low sample size		N	Low number of differences, likely to have arisen by chance
Risk of bias judgement (low, high, some concerns)	Low risk		Some concerns			Low risk	
EFFECT OF ASSIGNMENT TO INTERVENTION							
2.1 Were participants aware of their assigned intervention during the trial?	Y	Open-label	PY	Due to nature of the intervention (diet) Y			Non-blinded
2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	Open-label	PY	As above		Y	Non-blinded
2.3 If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?	PN	Few patients did not complete the study as per their allocated arm	PN	No suggestion that changes to the protocol occurred PN			A large proportion of patients in the glyburide arm switched to insulin, however this was reported to be consistent with rates observed in the literature
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA		NA			NA	

2.5 If Y/PY/NI to 2.4: Were these deviations from intended	NA	NA	NA	intervention balanced between groups?			
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Longitudinal follow-up, all patients with follow-up data were included	Y	All patients analysed as per their allocated arm	Y	The publication reports outcomes for the perprotocol population, but includes ITT analysis in the supplementary materials	
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the	NA	NA	NA	group to which they were randomized?			
Risk of bias judgement (low, high, some concerns)	Low risk	risk		Low	Low risk		

MISSING OUTCOME DATA						
3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Long-term follow-up rate was low and considered by the authors to be different from the initial cohort	Y	Outcome data was available for all participants	Y	Few participants with missing data reported
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?	N	No analysis carried out	NA		NA	
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	It is possible that patients in poorer health (e.g. due to maternal or neonatal complications) did not participate in long-term followup; however, lack of pregnancy outcomes for the full trial population prevents a comparison from being made with the pregnancy outcomes of participants with long-term follow-up data.	NA		NA	
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	See above	NA		NA	

Risk of bias judgement (low, high, some concerns)	High risk		Low risk	Low risk	
MEASUREMENT OF OUTCOME					
4.1 Was the method of measuring the outcome inappropriate?	PN	Due to nature of outcomes (e.g. mode of delivery and gestational age)	PY	Endpoints were not powered (due to small sample size)	PN	Due to nature of outcomes (e.g. mode of delivery and gestational age)
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	As above, due to nature of outcomes	PN	Due to nature of outcomes (e.g. gestational age)	PN	As above, due to nature of outcomes
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?	PN	Assessors were likely blind based on discussion in the paper	PN	Due to nature of outcomes (e.g. gestational age)	PY	Clinicians were not blinded to the group allocation
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PN	As above, due to nature of outcomes	NA		PN	As above, due to nature of outcomes
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?		NA		NA		NA

Risk of bias judgement (low, high, some concerns)	Low risk		High risk	Low risk	
SELECTION OF THE REPORTED RESULT					
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	Not specified in the publication for the outcomes of interest	Y	Endpoints described as pre-specified	Y	Endpoints described as prespecified
5.2. Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	Nature of outcomes means multiple possible definitions are unlikely	PN	Nature of outcomes means multiple possible definitions are unlikely	PN	Nature of outcomes means multiple possible definitions are unlikely

5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible analyses of the data?	N	Summary statistics used were identical for all variables of the same type (e.g. continuous or categorial)	N	Summary statistics used were identical for all variables of the same type (e.g. continuous or categorial)	N	Summary statistics used were identical for all variables of the same type (e.g. continuous or categorial)
Risk of bias judgement (low, high, some concerns)	Some concerns		Low risk		Low risk	
OVERALL RISK OF BIAS (low, high, some concerns)	High risk of bias		High risk of bias		Low risk	

Table 107: Cochrane RCTs

Question		Trout 2016		Reynolds 2017 (Study) 2	
RANDOMISATION PROCESS	**Answer**	**Notes**	**Answer**	**Notes**	
1.1 Was the allocation sequence random?	PY	No details of randomisation reported	Y		
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	No details of randomisation reported	Y		
1.3 Did the baseline differences between intervention groups suggest a problem with the randomization process?	PN	Few baseline characteristics reported, but appear balanced	PY	Only 23 women included in total, which makes it unfeasible to compare baseline characteristics. However, raw numbers appears balanced	
Risk of bias judgement (low, high, some concerns)	Some concerns		Low		
EFFECT OF ASSIGNMENT TO INTERVENTION					
2.1 Were participants aware of their assigned intervention during the trial?	PY	Assumed due to nature of intervention (diet)	Y		

2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	Assumed due to nature of intervention (diet)	Y		
2.3 If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?	NI	Unclear if deviations from the intended intervention occurred as the majority of participants did not complete food logs; however, any potential deviations are more likely due to non-adherence unrelated to study context	N	All women received the intervention as intended	
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA		NA		
2.5 If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA		NA		
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	ITT	Y	ITT	
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of they were randomized?	NA	MA	the failure to analyse participants in the group to which		
Risk of bias judgement (low, high, some concerns)	Some concerns		Low		
MISSING OUTCOME DATA					
3.1 Were data for this outcome available for all, or nearly all, participants	Change in maternal weight: 1 woman randomized?	Y	No missing outcome data reported	Y	missing data in the glibenclamide and 2 in the insulin groups
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?	NA		NA		
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA		NA		
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA		NA		

Risk of bias judgement (low, high, some concerns)	Low risk		Low	
MEASUREMENT F OUTCOME				
4.1 Was the method of measuring the outcome inappropriate?	PN	Due to nature of outcomes (e.g. mode of delivery and gestational age)	PN	
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	As above, due to nature of outcomes	PN	
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?	NI	Not specified	Y	
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PN	As above, due to nature of outcomes	PN	Outcomes are change in weight and mode of birth, bot unlikely to be measured differently between 2 groups
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA		NA	
Risk of bias judgement (low, high, some concerns)	Low risk		Low	
SELECTION OF THE R PORTED RESULT				
5.1 Were the data that produced this result analysed in accordance with a prespecified analysis plan that was finalized before unblinded outcome data were NI available for analysis?		Not specified	PY	Study was pre-registered on ClinicalTrial.gov and EudraCT
5.2. Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	Nature of outcomes means multiple possible definitions are unlikely	PN	
5.3 Is the numerical result being assessed likely to have been se basis of the results, from multiple eligible analyses of the data?	N	Summary statistics used were identical for all variables of the same type (e.g. continuous or categorial)	PN	

Risk of bias judgement (low, high, some concerns)	Some concerns	Low
OVERALL RISK OF BIAS (low, high, some concerns)	Some concerns	Low

Table 108: Cochrane RCTs

Question	Kokic 2018 1		Pellonpera 2016, Huhtala 2018 (Turku University Hospital) 2		Palatnik 2015, Casey 2015 (MFMU Network) 3	
RANDOMISATION PROCESS	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**

1.1 Was the allocation sequence random?	Y	Participants were randomised by block on clinicaltrials.gov	PY	Registered as a randomised clinical trial	PY	Computer-generated randomisation	Computer-generated randomisation
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	The process of allocation is controlled by a web-based computerized procedure	NI	No details provided	NI	No details provided	
1.3 Did the baseline differences between intervention groups suggest a problem with the randomization process?	N	No differences in baseline variables	N	No differences in baseline variables	N	No differences in baseline variables between overall treatment arms	
Risk of bias judgement (low, high, some concerns)	Low	–	Some concerns	–	Some concerns	–	
EFFECT OF ASSIGNMENT TO INTERVENTION							
2.1 Were participants aware of their assigned intervention during the trial?	Y	Due to the nature of the study, participants were not blinded	Y	Trial was open-label	Y	Blinding was not possible, due to practical differences between the treatment arms	
2.3 If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?	PN	There were no deviations from the intended intervention (exercise). Deviations from the dietary intervention were not recorded	N	No evidence of protocol changes	N	No evidence of protocol changes	
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA		NA		NA		

2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	N	Physicians and laboratory staff were blinded	Y	Trial was open-label	Y	Blinding was not possible, due to practical differences between the treatment arms	
2.5 If Y/PY/NI to 2.4: Were these deviations from intended	NA	NA	NA	intervention balanced between groups?			
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Analysis included all randomised patients except for those lost during pregnancy	Y	Analysis included all randomised patients with available outcomes data	Y	Analysis included all randomised patients with available outcomes data	
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in	NA	NA	NA	the group to which they were randomized?			
Risk of bias judgement (low, high, some concerns)	Low	–	Low	–	Low	–	
MISSING OUTCOME DATA							
3.1 Were data for this outcome available for all, or nearly all, participants randomized?	PY	Complete data available for >90% of randomised participants	PN	Complete data available for >85% of randomised participants, missing data was greater than the number of events for some dichotomous outcomes (microsomia)	Y	Complete data available for >95% of randomised participants	
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?	NA		NA		NA		
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA		NA		NA		
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the	NA	NA	NA	outcome depended on its true value?			
Risk of bias judgement (low, high, some concerns)	Low	–	Some concerns	–	Low	–	
MEASUREMENT OF OUTCOME							

4.1 Was the method of measuring the outcome inappropriate?	N	Outcomes were prespecified, with appropriate measurements and data collection	PN	Outcomes were prespecified, with appropriate measurements and data collection	N	Outcomes were prespecified, with appropriate measurements and data collection
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	Data was collected using the same measurement methods and frequency for both groups	N	Data was collected using the same measurement methods and frequency for both groups	N	Data was collected using the same measurement methods and frequency for both groups
4.3 If Y/PY/NI to 4.2: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA		PN	Outcomes were based on objective measurements or clinical record findings	PN	Outcomes were based on objective measurements or central review, where relevant
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA		PN	Outcomes were based on objective measurements or clinical record findings	PN	Outcomes were based on objective measurements or central review, where relevant
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA		NA		NA	
Risk of bias judgement (low, high, some concerns)	Low	–	Low	–	Low	–
SELECTION OF THE REPORTED RESULT						

4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?	N	All anthropometric measurements were performed by a blinded physiotherapist; laboratory staff who gathered data on maternal glycaemia	Y	Trial was open-label	NI	It is unclear whether all assessors were blinded
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Data were analysed according to a prespecified analysis plan, unclear whether this was finalised before outcome data were unblinded	N	Analyses were carried out as follow-up after the end of data collection and unclear whether data were blinded	N	Analyses were carried out as follow-up after the end of data collection and unclear whether data were blinded
5.2. Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	Results for all eligible measures are reported where outcomes were measured in different ways or at different timepoints	PN	Results for all eligible measures are reported where outcomes were measured in different ways or at different timepoints	PN	Outcomes were only measured one way at one timepoint. Justification was provided where data were not reported per treatment arm (e.g. shoulder dystocia).
5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible analyses of the data?	PN	Results for all analyses conducted are reported	PN	Results for all analyses conducted are reported	PN	Results for all analyses conducted are reported
OVERALL RISK OF BIAS (low, high, some concerns)		The study is judged to be at low risk of bias for all domains for this result.	Low / Some concerns	The study raises concerns in the domains of missing outcome data and selection of the reported results.	The study raises concerns in the domains of randomisation process, Some concerns and selection of the reported results.	
Risk of bias judgement (low, high, some concerns)	Low	–	Some concerns	–	Some concerns	–

Table 109. QUADAS-2 assessments for GDM screening studies

Question	Benhalima 2018b (BEDIP-N)		Iimura 2015		Khalafallah 2016		Saeedi 2018	
PARTICIPANT SELECTION	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**
Was a consecutive or random sample of participants enrolled? (Yes/No/Unclear)	Yes	Consecutive sample	Yes	Consecutive sample	Yes	Consecutive sample	Yes	Consecutive sample
Was a case-control design avoided? (Yes/No/Unclear)	Yes	Prospective cohort study	Yes	Prospective cohort study	Yes	Prospective cohort study	Yes	Cross-sectional study
Did the study avoid inappropriate exclusions? (Yes/No/Unclear)	Yes	Women with existing prediabetes or diabetes were excluded after initial FPG measurement (at 6– 14 weeks) but otherwise women with other risk factors were still included	Yes	No inappropriate exclusions	Unclear	No details reported on exclusions, but unlikely to have detrimental impact on results	Yes	No exclusions based on specific risk factors (based on the fact that risk factors were measured as part of the index test)
Could the selection of participants have introduced bias? (Low/High/Unclear)	Low	Based on signalling questions	Low	Based on signalling questions	Low	Based on signalling questions and judgement that inappropriate exclusions were unlikely	Low	Based on signalling questions
Is there concern that the included participants do not match the review question? (Low/High/Unclear)	Low	Singleton pregnancies in Belgium	Unclear	Singleton pregnancies in Japan. Population may be slightly different from a UK setting in terms of such as average BMI	Low	Singleton pregnancies in Australia	Low	Singleton pregnancies in Sweden
INDEX TESTS	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**

Were the index test results interpreted without knowledge of the reference standard? (Yes/No/Unclear)	Yes	GCT was administered before OGTT and was used as a deciding factor for going on to receive OGTT	Unclear	When measures were taken and interpreted is NR	Unclear	Measures were taken at the same time as the OGTT. Unclear if results were interpreted before the diagnosis was confirmed	Unclear	Risk factors were measured and recorded at first maternal visit (i.e. before the reference standard was received) but details of interpretation are NR
If a threshold was not prespecified because	No	Threshold for GCT was HbA1c were explored to see	Unclear	Threshold for makers NR were explored for	No	Arbitrary thresholds of	No/ N/A	Different thresholds used, was it pre-

REFERENCE STANDARD	Answer	Notes	Answer	Notes	Answer	Notes	Answer	Notes

Is the reference standard likely to correctly classify the test condition? (Yes/No/Unclear)	Yes	OGTT and WHO 2013 criteria	Yes	2-step screening using OGTT, with thresholds recommended by IADPSG and ADA	Yes	Although note that a 1-step test rather than a 2-step test is used	Unclear	A 1-step test rather than a 2-step test was used. The authors also note that a main limitation of the study was that when the material was collected, capillary whole blood sampling was used for OGTT and there is uncertainty about the conversion factors from capillary blood to venous plasma
Could the conduct or interpretation of the index test have introduced bias? (Low/High/Unclear)	High	Threshold was not prespecified	Unclear	Based on signalling questions	High	Unclear when measures taken and threshold not prespecified	High	Based on signalling questions
Is there concern that the index test, its								

	conduct, or interpretation differ from the review question? (Low/High/Unclear)	Low	lar to practice	High	Low	Low	Measurement of lipid/apolipoprotein biomarkers	currently recommended to detect GDM, the test is used	UK clinical practice those administered as currently recommends part of UK clinical
l t h o u g h H b A 1 c n o t S c r e e n i n g t e s t s i m i							is not part of current UK A	in standard practice to only offered to	screening based on risk but this is clinical practice assess blood glucose in factors women with risk factors diabetics
	Were the reference standard results interpreted without knowledge of the results of the index test? (Yes/No/Unclear)	Yes	Healthcare providers and participants were blinded with respect to results of the GCT, and all irrespective of the GCT result	Unclear	Details not reported	Unclear	Details not reported	Unclear	Details not reported	participants received an OGTT	
	Could the reference standard, its conduct, or its interpretation have introduced bias?	Low	Based on signalling questions	Low	Based on signalling questions	Low	Based on signalling questions	Unclear	Based on signalling questions		

(Low/High/Unclear)								
Is there concern that the test would receive the test in the UK (Low/High/Unclear)	Low	Same as UK definition in terms of thresholds, although note that only women with risk factors would receive the test in the UK	Low	Same as UK definition in terms of thresholds, although note that only reference standard recommends 1-step over 2-step	Low	Same as UK definition the target condition in the UK.	Same as UK definition in terms of thresholds, although note that only women with risk factors in the UK also would receive the test in the UK	Low
Did all participants receive a reference standard? (Yes/No Unclear)	Yes	All participants received OGTT, irrespective of their GCT result. Only participants who had received both GCT and OGTT were included in the analysis	Yes	All participants included in the analysis received the reference standard	Unclear	Assumed yes but not reported	No	All women were offered an OGTT but only 74% attended. Participants had significantly higher rates of GDM risk factors than those declining an OGTT, suggesting a greater risk of GDM compared to the total population. However, the authors do note that the prevalence difference should not substantially affect the test characteristics in view of the high overall attendance
Did participants receive the same reference standard? (Yes/No/Unclear)	Yes	All participants received OGTT. Methods were standardised across the study centres	Yes	All received OGTT	Unclear	Assumed yes but not reported	Yes	All received OGTT

Were all participants included in the analysis? (Yes/No/Unclear)	No	1.3% of participants were excluded from the analysis due to incomplete OGTT	No	Not all participants had biochemical or lipoprotein data (~50% missing)	Unclear	Assumed yes but not reported	Yes	All participants who received an OGTT were included in the analysis
		2-step						
PARTICIPANT FLOW	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**	**Answer**	**Notes**
Was there an appropriate interval between the index test(s) and the reference standard? (Yes/No/Unclear)	Yes	OGTT was given before any treatment decisions were made	Unclear	Timepoint of measurement of lipid/apolipoprotein markers NR	Yes	OGTT and HbA1c were measured at the same time, before any treatment decisions were made	Yes	OGTT was given before any treatment decisions were made

Could the participant flow have introduced	Low	Even though not all participants were included	Unclear	Based on signalling questions	Based on signalling	Unclear	Low	Based on signalling questions and

bias? (Low/High/Unclear)	in the analysis, this is not judged to be a concern for risk of bias, as the number excluded due to not having a complete OGTT was very small (1.3%)			questions		assumption that overall attendance was high enough to not impact the test characteristics

Table 110: QUADAS-2 assessments for GDM screening studies

Question	Temming 2016		Ohara 2016		Ryser Ruetschi 2016	
PARTICIPANT SELECTION	Answer	Notes	Answer	Notes	Answer	Notes
Was a consecutive or random sample of participants enrolled? (Yes/No/Unclear)	Yes	All consecutive patients undergoing the relevant test	Yes	All women without hyperemesis gravidarum who delivered within the specified period	Yes	Tests included in the analysis were consecutive and not selected
Was a case-control design avoided? (Yes/No/Unclear)	Yes	Retrospective cohort study	Yes	Retrospective cohort study, a case-control design was only applied for a different question within the same study	Yes	Cross-sectional study
Did the study avoid inappropriate exclusions? (Yes/No/Unclear)	Yes	Exclusions were appropriate and unlikely to lead to bias	Unclear	Women with hyperemesis gravidarum were analysed as part of a separate study group, all other exclusions appropriate (women with prior diabetes)	Yes	No exclusion criteria were applied
Could the selection of participants have introduced bias? (Low/High/Unclear)	Low	Overall low risk of bias introduced by the selection of participants	Unclear	Unclear risk of bias introduced by the selection of participants	Low	Overall low risk of bias introduced by the selection of participants

Is there concern that the included participants do not match the review question? (Low/High/Unclear)	Unclear	Overall low-risk pregnancies representative of the screening population; however, ~50% of patients were African American, so ethnic similarities to UK population unclear	Unclear	Overall low-risk pregnancies representative of the screening population; however, women with hyperemesis gravidarum were analysed as part of a separate study group and the effect of this is unclear	Low	Overall low-risk pregnancies representative of the screening population	

INDEX TESTS	Answer	Notes	Answer	Notes	Answer	Notes
Were the index test results interpreted without knowledge of the reference standard? (Yes/No/Unclear)	Yes	Screening results were interpreted before the diagnostic test was	Yes	Screening results were interpreted before the diagnostic test was	Unclear	Unclear if screening results were interpreted without knowledge of the reference
		administered and diagnosis was confirmed		administered and diagnosis was confirmed		standard
If a threshold was used, was it pre-specified? (Yes/No/Unclear)	Yes	Threshold pre-specified according to published guidelines, data also analysed in categorical increments	Yes	Threshold was pre-specified	Yes	Thresholds pre-specified according to published guidelines
Could the conduct or interpretation of the index test have introduced bias? (Low/High/Unclear)	Low	Overall low risk of bias introduced by the conduct or interpretation of the index test; it is likely that staff conducting the test could have had foreknowledge of who was at risk, however, this is unlikely to influence the outcome of the test	Low	Overall low risk of bias introduced by the conduct or interpretation of the index test; it is likely that staff conducting the test could have had foreknowledge of who was at risk, however, this is unlikely to influence the outcome of the test	Low	Overall low risk of bias introduced by the conduct or interpretation of the index test; it is likely that staff conducting the test could have had foreknowledge of the reference standard results, however, this is unlikely to influence the outcome of the test
Is there concern that the index test, its conduct, or interpretation differ from the review question? (Low/High/Unclear)	Low	Screening test (50 g GCT) is similar to screening tests administered as part of UK clinical practice	Low	Screening test (50 g GCT) is similar to screening tests administered as part of UK clinical practice	Low	Screening test (FPG) is similar to screening tests administered as part of UK clinical practice
REFERENCE STANDARD	Answer	Notes	Answer	Notes	Answer	Notes

Is the reference standard likely to correctly classify the test condition? (Yes/No/Unclear)	Yes	GDM confirmed by 100 g OGTT at ≥24 weeks of gestation according to NDDG criteria and CC criteria	Yes	GDM confirmed by 75 g OGTT at ≥24 weeks of gestation, on the basis of the universal criteria established by the IADPSG	Yes	GDM confirmed by 75 g OGTT at ≥24 weeks of gestation, on the basis of the universal criteria established by the IADPSG
Were the reference standard results interpreted without knowledge of the results of the index test? (Yes/No/Unclear)	No	Screening results were likely known by the investigator making the final diagnosis	No	Screening results were known by the investigator making the final diagnosis	No	Screening results were likely known by the investigator making the final diagnosis

Could the reference standard, its conduct, or its interpretation have introduced bias? (Low/High/Unclear)	Low	Overall low risk of bias introduced, despite investigator awareness of screening results, as objective diagnostic criteria applied	Low	Overall low risk of bias introduced, despite investigator awareness of screening results, as objective diagnostic criteria applied	Low	Overall low risk of bias introduced, despite investigator awareness of screening results, as objective diagnostic criteria applied
Is there concern that the target condition as defined by the reference standard does not match the review question? (Low/High/Unclear)	Low	The reference standard (100 g OGTT) defined GDM using NDDG or CC criteria, which is similar to current UK clinical practice	Low	The reference standard (75 g OGTT) defined GDM using IADPSG, which is similar to current UK clinical practice	Low	The reference standard (75 g OGTT) defined GDM using IADPSG, which is similar to current UK clinical practice
PARTICIPANT FLOW	Answer	Notes	Answer	Notes	Answer	Notes
Was there an appropriate interval between the index test(s) and the reference standard? (Yes/No/Unclear)	Yes	Index test likely administered ~1 week before reference standard	Unclear	Exact interval unclear; reference standard was likely given before any change in clinical care occurred	Yes	Index test (FPG) administered at the star of reference standard test
Did all participants receive a reference standard? (Yes/No Unclear)	No	Participants with a normal index test result (50 g GCT <140 mg/dL) did not receive the reference standard Not all participants with an elevated index test result (50 g GCT ≥140 mg/dL) received the reference standard	No	Only participants who tested positive on the screening test (50 g GCT) were given the reference standard	Yes	All participants for whom data was collected received the reference standard test
Did participants receive the same reference standard? (Yes/No/Unclear)	Yes	All participants who received a reference standard received the same test	Yes	All participants who received a reference standard received the same test	Yes	All participants received the same reference standard test

Were all participants included in the analysis? (Yes/No/Unclear)	No	Participants with a normal index test result and participants who did not receive the reference standard were not included in the analysis	No	All women who were eligible were included in the analysis; participants with hyperemesis gravidarum were analysed as a separate group	No	1.2% of participants were excluded due to incomplete data, all other participants included in the analysis	
Could the participant flow have introduced bias? (Yes/No/Unclear)	Yes	A significant proportion of women were removed from the analysis	Yes	Not all participants received the reference standard test and participants with hyperemesis gravidarum were analysed as a separate group	No	Women who underwent the index test were all equally likely to develop and be diagnosed with gestational diabetes in the same manner; excluded participants unlikely to influence risk of bias	

Question	**Pawelec 2009**		**Maesa 2018**		**Kosus 2012**	
PARTICIPANT SELECTION	Answer	Notes	Answer	Notes	Answer	Notes

Table 111: QUADAS-2 assessments for GDM screening studies

Was a consecutive or random sample of participants enrolled? (Yes/No/Unclear)	Unclear	Unclear how patients were enrolled	Yes	All women with complete data in a setting where screening is performed on all pregnant women	Unclear	Unclear how patients were enrolled
Was a case-control design avoided? (Yes/No/Unclear)	Yes	Prospective cohort study	Yes	Retrospective cohort study	Yes	Retrospective cohort study
Did the study avoid inappropriate exclusions? (Yes/No/Unclear)	Yes	No patients were excluded	Yes	No patients were excluded	No	Patients with potential previous diabetic pregnancy were excluded, which might not be appropriate for the study aim
Could the selection of participants have introduced bias? (Low/High/Unclear)	Unclear	Overall risk of bias unknown, as patient enrolment procedures unclear	Low	Overall low risk of bias introduced by the selection of participants	High	Overall risk of bias high due to inappropriate patient exclusions
Is there concern that the included participants do not match the review question? (Low/High/Unclear)	Unclear	No information is provided on the risk status of patients, including exclusion based on multiple pregnancies	Low	Overall low-risk pregnancies representative of the screening population	Unclear	Unclear if the pregnancies included in this study are representative of the screening population
INDEX TESTS	Answer	Notes	Answer	Notes	Answer	Notes
Were the index test results interpreted without knowledge of the reference standard? (Yes/No/Unclear)	Yes	Screening results were interpreted before the diagnostic test was administered and diagnosis was confirmed	Unclear	Unclear if index test screening results were interpreted before the diagnosis was confirmed	No	Index test screening results were interpreted with knowledge of the diagnosis
If a threshold was used, was it pre-specified? (Yes/No/Unclear)	Yes	Threshold pre-specified according to published guidelines	No	Test accuracy was reported for a range of different thresholds	No	Thresholds were developed on the basis of the results of the reference standard

Could the conduct or interpretation of the index test have introduced bias? (Low/High/Unclear)	Low	Overall low risk of bias introduced by the conduct or interpretation of the index test; it is likely that staff conducting the test could have had foreknowledge of who was at risk, however, this is unlikely to influence the outcome of the test	High	Overall high risk of bias introduced by the conduct and interpretation of the index test due to lack of pre-specified thresholds and unclear index test result interpretation timeframe	High	Overall high risk of bias introduced by the conduct and interpretation of the index test due to lack of 'blinding' in index test interpretation and lack of pre-specified thresholds
Is there concern that the index test, its conduct, or interpretation differ from the review question? (Low/High/Unclear)	Low	Screening test (50 g GCT by capillary blood sampling) is similar to screening tests administered as part of UK clinical practice	Low	Screening test (fasting glycaemia) is similar to screening tests administered as part of UK clinical practice	Low	Unclear if the screening test strategy includes GCT prior to OGTT, but each test is commonly used in UK clinical practice
REFERENCE STANDARD	Answer	Notes	Answer	Notes	Answer	Notes
Is the reference standard likely to correctly classify the test condition? (Yes/No/Unclear)	Yes	GDM confirmed by 75 g OGTT at ≥24 weeks of gestation according to Carpenter and Coustan International Workshop−Conference on	Yes	GDM confirmed by 100 g OGTT ≥24 weeks of gestation according to NDDG criteria	Yes	GDM confirmed by 100 g OGTT between 24 and 28 weeks of gestation according to CC criteria
		GDM criteria				
Were the reference standard results interpreted without knowledge of the results of the index test? (Yes/No/Unclear)	No	Screening results were likely known by the investigator making the final diagnosis	No	Screening results were likely known by the investigator making the final diagnosis	Unclear	Unclear if index test cut-off values and results were known by the investigator making the final diagnosis
Could the reference standard, its conduct, or its interpretation have introduced bias? (Low/High/Unclear)	Low	Overall low risk of bias introduced, despite investigator awareness of screening results, as objective diagnostic criteria applied	Low	Overall low risk of bias introduced, despite investigator awareness of screening results, as objective diagnostic criteria applied	Low	Overall low risk of bias introduced, despite investigator awareness of screening results, as objective diagnostic criteria applied
Is there concern that the target condition as defined by the reference standard does not match the review question? (Low/High/Unclear)	Low	The reference standard (75 g OGTT) defined GDM using CC criteria, which is similar to current UK clinical practice	Low	The reference standard (100 g OGTT) defined GDM using NDDG criteria, which is similar to current UK clinical practice	Low	The reference standard (100 g OGTT) defined GDM using C&C criteria, which is similar to current UK clinical practice
PARTICIPANT FLOW	Answer	Notes	Answer	Notes	Answer	Notes

Was there an appropriate interval between the index test(s) and the reference standard? (Yes/No/Unclear)	Yes	The reference standard was given before any change in clinical care occurred	Unclear	Exact interval unclear; reference standard was likely given before any change in clinical care occurred	Yes	The reference standard was given before any change in clinical care occurred
Did all participants receive a reference standard? (Yes/No Unclear)	No	Only participants who tested positive on the screening test (50 g GCT by venous or finger capillary blood) were given the reference standard	No	Only participants who tested positive on the screening test (50 g GCT) were given the reference standard	No	Only participants who tested positive on the screening test (50 g GCT) were given the reference standard
Did participants receive the same reference standard? (Yes/No/Unclear)	Yes	All participants who received a reference standard received the same test	Yes	All participants who received a reference standard received the same test	Yes	All participants who received a reference standard received the same test
Were all participants included in the analysis? (Yes/No/Unclear)	Unclear	Only number of participants included in analysis was reported, unclear if representative of all patients screened	Yes	All screened women were included in the final analysis	Unclear	Only number of participants included in analysis was reported, unclear if representative of all patients screened
Could the participant flow have introduced bias? (Yes/No/Unclear)	Yes	Only participants who tested positive on screening were given the reference standard test, unclear if all screened women included in analysis	Yes	Only participants who tested positive on screening were given the reference standard test, time interval between index test and reference standard unclear	Yes	Only participants who tested positive on screening were given the reference standard test, unclear if all screened women included in analysis

Table 112: QUADAS-2 assessments for GDM screening studies

Question	Project Viva, Gingras 2018	
PARTICIPANT SELECTION	**Answer**	**Notes**
Was a consecutive or random sample of participants enrolled? (Yes/No/Unclear)	Yes	Consecutive sample
Was a case-control design avoided? (Yes/No/Unclear)	Yes	Cohort study
Is the reference standard likely to correctly classify the test condition? (Yes/No/Unclear)	Yes	2-step test, consistent with some major recommendations
Were the reference standard results interpreted without knowledge of the results of the index test? (Yes/No/Unclear)	Yes	OGTT was performed several years before fructosamine levels were measured

Did the study avoid inappropriate exclusions? (Yes/No/Unclear)	Yes	Details of exclusions were given and included missing fructosamine measurement (the majority); samples with gross or moderate haemolysis/lipaemia; pregestational type 1 or type 2 diabetes; missing glucose tolerance assessment in pregnancy; missing covariates. Exclusions are judged to be appropriate and furthermore, the authors carried out a sensitivity analysis using inverse probability weighting (IPW) to assess the impact of the missing fructosamine measurements and found no difference in the AUC value for test accuracy
Could the selection of participants have introduced bias? (Low/High/Unclear)	Low	Based on signalling questions
Is there concern that the included participants do not match the review question? (Low/High/Unclear)	Low	Singleton live births in USA setting

INDEX TESTS	Answer	Notes
Were the index test results interpreted without knowledge of the reference standard? (Yes/No/Unclear)	Unclear	Samples and blood glucose measurements were in 1999–2002 but fructosamine was measured from frozen samples in 2016–2017. Details on whether the assessors were blinded to the glucose results is NR
If a threshold was used, was it pre-specified? (Yes/No/Unclear)	Yes	Three thresholds for fructosamine were prespecified in the methods
Could the conduct or interpretation of the index test have introduced bias? (Low/High/Unclear)	Unclear	Based on signalling questions
Is there concern that the index test, its conduct, or interpretation differ from the review question? (Low/High/Unclear)	High	Fructosamine measurement is not currently recommended in UK clinical practice

REFERENCE STANDARD	Answer	Notes
Could the reference standard, its conduct, or its interpretation have introduced bias?		
Is there concern that the target condition as defined by the reference standard does not match the review question? (Low/High/Unclear)	Low	Same as UK definition in terms of thresholds, although note that only women with risk factors would receive the test in the UK. UK also recommends 1-step over 2-step

PARTICIPANT FLOW	Answer	Notes
(Low/High/Unclear)	Low	Based on signalling questions

Was there an appropriate interval between the index test(s) and the reference standard? (Yes/No/Unclear)	Yes	Blood sample for fructosamine was collected at the same time as OGTT, i.e. before treatment decisions (even though the measurement of fructosamine was not performed until years later)
Did all participants receive a reference standard? (Yes/No Unclear)	No	Participants who did not receive assessment of glucose levels (N=21) were not included in the analysis
Did participants receive the same reference standard? (Yes/No/Unclear)	Yes	
Were all participants included in the analysis? (Yes/No/Unclear)	No	However judged unlikely to cause an issue
Could the participant flow have introduced bias? (Low/High/Unclear)	Low	Based on signalling questions

Table 113: PROBAST assessment of screening studies with models

Question	van Leeuwen 2010		Theriault 2014	
TYPE of PREDICTION STUDY	**Answer**	**Notes**	**Answer**	**Notes**
Classify the evaluation based on its aim (i.e. what is the type of development (multiple only/Development and	Development only	Prediction model previously developed in other participants	Validation only	External validation of four models prediction study)? (Development logistic regression) without external validation
PARTICIPANTS	**Answer**	**Notes**	**Answer**	**Notes**

Risk of Bias	1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data? (Y/PY/PN/N/NI)	PN		Data from an existing prospective cohort study; however, collection of data (baseline characteristics, glucose measurements and medical history) unlikely to be unsuitable for the purposes of creating a model	Y		Data are from a prospective longitudinal cohort
	1.2 Were all inclusions and exclusions of participants appropriate? (Y/PY/PN/N/NI)	PY		Inclusion/exclusion criteria were appropriate (only women with known pregestational diabetes mellitus and women who were first seen after 20 weeks of gestation were excluded from the study) Inclusion/exclusion criteria were appropriate (only women with known pregestational diabetes mellitus, multiple pregnancy, uncertain diagnosis and delivery outside of study centres)	Y		

					Applicability	What is the concern that the definition, assessment	Low	assessment and	
						or timing of predictors in the model do not match			timi
						What is the risk of bias introduced by selection of participants (Low/High/Unclear)	Low		Ove
						collection of data in the cohort study was unlikely to be unsuitable for the purposes of creating a model and appropriate inclusion/exclusion criteria			were
Applicability	What is the concern that the included participants and setting do not match the review question? (Low/High/Unclear)	Low			Included participants, the selection criteria used and the setting used in the primary study are relevant to the review question		Low		
2.2 Were predictor assessments made without knowledge of outcome data? (Y/PY/PN/N/NI)		Y			Predictor assessments (baseline characteristics) were carried out at intake and before any outcome (blood glucose) assessment		PY		

er	Notes
	Included participants, the selection criteria used and the setting used in the primary study are relevant to the review question

	Definitions of predictors (health risk factors) and their assessment were the same for all participants				
	Predictors (health risk factors) were likely assessed before knowledge of the outcome; most predictors are objective measures				
Risk of Bias	2.3 Are all predictors available at the time the model is intended to be used? (Y/PY/PN/N/NI)	Y	All included predictors (baseline characteristics/medical history) would be available at the time the model is intended to be used for prediction	Y	All included predictors (health risk factors/anthropometric measurements) would be available at the time the models are intended to be used for prediction
	What is the risk of bias introduced by predictors or the review question? (Low/High/Unclear)	Low	Low risk of bias introduced by predictors or their assessment model match the review question	Low	Low risk of bias introduced by predictors or their assessment model match the review question
OUTCOME		**Answer**	**Notes**	**Answer**	**Notes**
	3.1 Was the outcome determined appropriately? (Y/PY/PN/N/NI)	Y	determination used (blood glucose or diagnosis of GDM) is considered acceptable by the WHO patients the outcome was determined by use of insulin during pregnancy	PY	The method of outcome determination used (blood glucose) is considered acceptable according to the 2008 Canadian Diabetes Association recommendations; in some 1999 guidelines on the topic
	3.2 Was a pre-specified or standard outcome definition used? (Y/PY/PN/N/NI)	Y	Method of outcome determination was objective (blood glucose measurement)	PY	Methods of outcome determination were likely objective (blood glucose measurement or use of insulin during pregnancy)
Risk of Bias	3.3 Were predictors excluded from the outcome definition? (Y/PY/PN/N/NI)	Y	None of the predictors outcome definition (blood glucose or GDM)	PY	None of the predictors definition (diagnosis of GDM by blood glucose or insulin

		(health				
		risk				
(baseline factors/anthropometric characteristics/medical measurements) were history) were included in the Y included in the outcome use during pregnancy)						
3.4 Was the outcome defined and determined in a similar way for all participants? (Y/PY/PN/N/NI)	Y		Outcomes were defined and recorded, as predictor determined using (baseline established blood glucose characteristics/medical history) testing methodology was not timedependent and (outcome was assessed at the random plasma glucose or GCT, followed by OGTT if appropriate time (at 24–28 positive) for all participants; weeks of only data from one centre gestation) used	PN	outcome determination The outcome was defined (using established bloodiagnosis of GDM by blood d glucose or insulin use during glucose testing methodology pregnancy) was likely (random plasma glucose or appropriate to enable the GCT, followed by OGTT if correct type and positive) in some patients representative number of and by use of insulin during relevant outcomes to be pregnancy in others recorded	relevant outcomes to be
			measurements) and			
3.5 Was the outcome determined without knowledge of predictor information? (Y/PY/PN/N/NI)	NI		Predictor information might available, but unclear if known by assessors determining the outcome status	Predictor information might have been available, but	NI	have been unclear if known by assessors determining the outcome status

	3.6 Was the time interval between predictor assessment and outcome determination appropriate? Y (Y/PY/PN/N/NI)			No information on time interval needed to enable the correct type and representative number of	PY	Time interval between predictor assessment (health risk factors/anthropometric
	What is the risk of bias introduced by the outcome by the use of different methodologies	Overall high risk of bias introduced Low	or its determination? (Low/High/Unclear)			
	Overall low risk of bias or its determination	to define and determine outcome	introduced by the outcome	High		
pplicability	What is the concern that the outcome, its definition, timing or determination do not match the review question? (Low/High/Unclear)	Low		Definition, timing and determination of outcome match the review question	High	Definition, timing and determination of the outcome by blood glucose measurement matches the review question, but definition/determination of the outcome by insulin use in pregnancy does not match the review question
NALYSIS		**Answer**	**Notes**		**Answer**	**Notes**

	Number of participants with				Number of participants with ~~outcome (GDM) is 381 and~~ there are 8 candidate	
Risk of Bias						4.2 Were continuous and appropriately? (Y/PY/PN/
					predictor parameters across	
4.1 Were there a reasonable number of participants	N	there are 8 candidate	Y with the outcome? (Y/PY/PN/N/NI)		outcome (GDM) is 24 and	
				EPV =3	predictor parameters, so	

4.3 Were all enrolled participants included in the analysis? (Y/PY/PN/N/NI)	PN	PY	Only women from 1 of the 2 study centres were included; participants (7208/7929) included in the analysis		Over 90% of enrolled participants unclear number of enrolled were
4.4 Were participants with missing data handled appropriately? (Y/PY/PN/N/NI)	PY		Missing results of OGTT in all women were handled using multiple imputation; unclear how missing values for the predictors were handled	PN	Participants with absence of screening and/or diagnostic tests and gestational age at delivery unknown or before 32 weeks of gestation were excluded from the analysis; however, data on participants who did not undergo screening and/or diagnostic tests, but delivered after 32 weeks of gestation was extracted from the medical records and still included in the analysis (outcome was defined by use of insulin in pregnancy)
Model development studies only			Predictors initially selected on the basis of existing knowledge, significance		
4.5 Was selection of predictors based on univariable analysis avoided? (Y/PY/PN/N/NI)	Y		threshold of univariable analysis adjusted to 30% to avoid the erroneous exclusion of a potential relevant predictive variable	N/A	N/A
4.6 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately? (Y/PY/PN/N/NI)	PY		No relevant complexities in the data to account for	PY	No relevant complexities in the data to account for

			Discrimination (area under the receiver operating characteristic curve) evaluated appropriately; however, calibration only assessed using goodness-		4.7 Were relevant model performance measures of-fit test (Hosmer-
	(Y/PY/PN/N/NI)	Lemeshow) and			N PY evaluated appropriately?
			classification measures (sensitivity, specificity) presented using arbitrary		
	Model development studies only 4.8 Were model overfitting and optimism in model performance accounted for? (Y/PY/PN/N/NI)	PN	Unclear if internal model validation was performed; all model parameters were uniformly shrunken to adjust for optimism	N/A	N/A
			predicted probability thresholds, which were not pre-specified		
	Calibration plots available (supplementary figures) and discrimination (area under the receiver operating characteristic curve) evaluated appropriately; classification measures (sensitivity, specificity, predictive values) were presented				
bias/Unclear risk of bias) Overall high risk of bias	What is the risk of bias introduced by the analysis? (Low/High/Unclear)		Risk of bias may have been introduced by the analysis, particularly by the low number High		Risk of bias may have been introduced by the analysis, particularly by the handling
Overall high risk of bias			High	of participants with the outcome	of missing data
	OVERALL ASSESSMENT	Answer	Notes	Answer	Notes
	Overall applicability judgement (Low concerns for applicability/High concerns for applicability/Unclear concerns for applicability)	Low concerns for applicability	Low concerns for applicability across all domains	High concerns for applicability	High concerns for applicability, due to outcome definition/assessment
	using predicted probability thresholds maximising the Youden index, likely derived from the data set at hand **Model development studies only**				
	4.9 Do predictors and their assigned weights in the PY final model correspond to the results from the reported multivariable analysis? (Y/PY/PN/N/NI) Predictors and regression coefficients in the final model fully reported; regression				
	coefficient results from the multivariable analysis unclear, so comparison cannot be made	N/A	N/A		
Overall risk of bias judgement (Low risk of bias/High risk of					assessment and model

High

introduced analysis

by the outcome introduced by the model High analysis

Appendix 6 – UK NSC reporting checklist for evidence summaries

All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table 114.

Table 114. UK NSC reporting checklist for evidence summaries

	Section	Item	Page no.
1.	TITLE AND SUMMARIES		
1.1	Title sheet	Identify the review as a UK NSC evidence summary.	Title page
1.2	Plain English summary	Plain English description of the executive summary.	
1.3	Executive summary	Structured overview of the whole report. To include: the purpose/aim of the review; background; previous recommendations; findings and gaps in the evidence; recommendations on the screening that can or cannot be made on the basis of the review.	
2.	INTRODUCTION AND APPROACH		
2.1	Background and objectives	Background – Current policy context and rationale for the current review – for example, reference to details of previous reviews, basis for current recommendation, recommendations made, gaps identified, drivers for new reviews Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary, criteria they address, and number of studies included per question, description of the overall results of the literature search.	

		Method – briefly outline the rapid review methods used.		
2.2	Eligibility for inclusion in the review	State all criteria for inclusion and exclusion of studies to the review clearly (PICO, dates, language, study type, publication type, publication status etc.) To be decided *a priori*.		
2.3	Appraisal for quality/risk of bias tool	Details of tool/checklist used to assess quality, e.g. AMSTAR.	QUADAS 2, CASP, SIGN,	
3.	SEARCH STRATEGY AND STUDY SELECTION (FOR EACH KEY QUESTION)			
3.1	Databases/ final search.	Give details of all databases searched (including	sources	platform/interface and coverage dates) and date of searched
3.2	Search strategy and results	Present the full search strategy for at least one database (usually a version of Medline), including limits and search filters if used. Provide details of the total number of (results from each database searched), number of duplicates removed, and the final number of unique records to consider for inclusion.		
3.3	Study selection	State the process for selecting studies – inclusion and	140 exclusion criteria, number of reviewers, any	
		number of studies screened by title/abstract and full text, cross checking carried out.		
4.	STUDY LEVEL REPORTING OF RESULTS (FOR EACH KEY QUESTION)			
5.	QUESTION LEVEL SYNTHESIS			
4.1	Study level reporting, results and risk of bias assessment	For each study, produce a table that includes the full citation and a summary of the data relevant to the question (for example, study size, PICO, follow-up period, outcomes reported, statistical analyses etc.). Provide a simple summary of key measures, effect estimates and confidence intervals for each study where available. For each study, present the results of any assessment of quality/risk of bias.	Study level reporting: Quality assessment:	

5.1	Description of studies	For each question, give numbers of studies screened, assessed for eligibility, and included in the review, with summary reasons for exclusion.
5.2	Combining and presenting the evidence	Provide a balanced discussion of the body of evidence which avoids over reliance on one study or set of findings. Consideration of four components should inform the reviewer's judgement on whether the criterion is 'met', 'not met' or 'uncertain': quantity; quality; applicability and consistency.
5.3	Summary of findings	Provide a description of the evidence reviewed and included for each question, with reference to their eligibility for inclusion. Summarise the main findings including the quality/risk of bias issues for each question. Have the criteria addressed been 'met', 'not met' or 'uncertain'?
6.	REVIEW SUMMARY	
6.1	Conclusions and policy implications	Do findings indicate whether screening should be recommended? Is further work warranted? Are there gaps in the evidence highlighted by the review?
6.2	Limitations	Discuss limitations of the available evidence and of the review methodology if relevant.

References

1. Jacklin PB, Maresh MJA, Patterson CC, et al. A cost-effectiveness comparison of the NICE 2015 and WHO 2013 diagnostic criteria for women with gestational diabetes with and without risk factors. BMJ Open 2017;7:e016621.
2. National Institute for Health and Care Excellence (NICE). Diabetes in pregnancy: management from preconception to the postnatal period (NG3). Available at: https://www.nice.org.uk/guidance/ng3, 2015.
3. HAPO Study Cooperative Research Group. Hyperglycemia and adverse pregnancy outcomes. New England Journal of Medicine 2008;358:1991-2002.
4. Farrar D, Simmonds M, Griffin S, et al. The identification and treatment of women with hyperglycaemia in pregnancy: An analysis of individual participant data, systematic reviews, meta-analyses and an economic evaluation. Health Technology Assessment 2016;20:1-382.
5. Waugh N, Royle P, Clar C, et al. Screening for hyperglycaemia in pregnancy: a rapid update for the National Screening Committee. Health Technology Assessment 2010;14:1-183.
6. NHS. Overview Gestational Diabetes, 2019. Available at: https://www.nhs.uk/conditions/gestational-diabetes/.
7. Plows JF, Stanley JL, Baker PN, et al. The Pathophysiology of Gestational Diabetes Mellitus. International Journal of Molecular Sciences 2018;19.
8. Wendland EM, Torloni MR, Falavigna M, et al. Gestational diabetes and pregnancy outcomes-a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic criteria. BMC pregnancy and childbirth 2012;12:23.
9. Bellamy L, Casas J-P, Hingorani AD, et al. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. The Lancet 2009;373:1773-1779.
10. Shah BR, Retnakaran R, Booth GL. Increased risk of cardiovascular disease in young women following gestational diabetes mellitus. Diabetes care 2008;31:1668-1669.
11. Boney CM, Verma A, Tucker R, et al. Metabolic syndrome in childhood: association with birth weight, maternal obesity, and gestational diabetes mellitus. Pediatrics 2005;115:e290-e296.
12. Lawlor D, Fraser A, Lindsay R, et al. Association of existing diabetes, gestational diabetes and glycosuria in pregnancy with macrosomia and offspring body mass index, waist and fat mass in later childhood: findings from a prospective pregnancy cohort. Diabetologia 2010;53:89-97.
13. Correa PJ, Vargas JF, Sen S, et al. Prediction of gestational diabetes early in pregnancy: targeting the long-term complications. Gynecologic and Obstetric Investigation 2014;77:145-149.
14. Makgoba M, Savvidou MD, Steer PJ. An analysis of the interrelationship between maternal age, body mass index and racial origin in the development of gestational diabetes mellitus. British Journal of Obstetrics and Gynaecology 2012;119:276-82.

15. Sacks DA, Hadden DR, Maresh M, et al. Frequency of gestational diabetes mellitus at collaborating centers based on IADPSG consensus panel-recommended criteria: the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study. Diabetes Care 2012;35:526-8.
16. WHO. Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Part 1: Diagnosis and Classification of Diabetes Mellitus. 1999.
17. WHO. WHO Expert Committee on Diabetes Mellitus: second report. World Health Organization Technical Report Series 1980;646:1-80.
18. Duran A, Saenz S, Torrejon MJ, et al. Introduction of IADPSG criteria for the screening and diagnosis of gestational diabetes mellitus results in improved pregnancy outcomes at a lower cost in a large cohort of pregnant women: the St. Carlos Gestational Diabetes Study. Diabetes Care 2014;37:2442-50.
19. Rani PR BJ. Screening and Diagnosis of Gestational Diabetes Mellitus, Where Do We Stand. Journal of Clinical and Diagnostic Research 2016;10:QE01–QE04.
20. Yuen L, Wong VW, Simmons D. Ethnic disparities in gestational diabetes. Current diabetes reports 2018;18:68.
21. Dornhorst A, Paterson CM, Nicholls JS, et al. High prevalence of gestational diabetes in women from ethnic minority groups. Diabetic Medicine 1992;9:820-5.
22. Torloni MR, Betran AP, Horta BL, et al. Prepregnancy BMI and the risk of gestational diabetes: a systematic review of the literature with meta-analysis. Obesity Reviews 2009;10:194-203.
23. Najafi F, Hasani J, Izadi N, et al. The effect of prepregnancy body mass index on the risk of gestational diabetes mellitus: A systematic review and dose-response meta-analysis. Obesity Reviews 2019;20:472-486.
24. Anderwald C, Tura A, Winhofer Y, et al. Glucose Absorption in Gestational Diabetes Mellitus During an Oral Glucose Tolerance Test. Diabetes Care 2011;34:1475.
25. Paradisi G, Biaggi A, Ferrazzani S, et al. Abnormal Carbohydrate Metabolism During Pregnancy. Diabetes Care 2002;25:560.
26. ADA. Standards of Medical Care in Diabetes—2014. Diabetes Care 2014;37:S14.
27. O'Sullivan JB, Mahan CM. CRITERIA FOR THE ORAL GLUCOSE TOLERANCE TEST IN PREGNANCY. Diabetes 1964;13:278-85.
28. National Diabetes Data Group. Classification and Diagnosis of Diabetes Mellitus and Other Categories of Glucose Intolerance. Diabetes 1979;28:1039.
29. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. American Journal of Obstetrics and Gynecology 1982;144:768-73.

30. ADA. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 1997;20:1183-97.
31. WHO. Diabetes mellitus. Report of a WHO Study Group. World Health Organization Technical Report Series 1985;727:1-113.
32. International Association of Diabetes and Pregnancy Study Groups (IADPSG) Consensus Panel. International Association of Diabetes and Pregnancy Study Groups Recommendations on the Diagnosis and Classification of Hyperglycemia in Pregnancy. Diabetes Care 2010;33:676.
33. ADA. Standards of Medical Care in Diabetes—2011. Diabetes Care 2011;34:S11.
34. WHO. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy. Available at: https://www.who.int/diabetes/publications/Hyperglycaemia_In_Pregnancy/en/, 2013.
35. Blumer I, Hadar E, Hadden DR, et al. Diabetes and pregnancy: an endocrine society clinical practice guideline. Journal of Clinical Endocrinology and Metabolism 2013;98:4227-49.
36. Vandorsten JP, Dodson WC, Espeland MA, et al. NIH consensus development conference: diagnosing gestational diabetes mellitus. NIH State-of-the-Science Consensus Statement 2013;29:1-31.
37. Sacks DB. Diagnosis of Gestational Diabetes Mellitus: It Is Time for International Consensus. Clinical Chemistry 2014;60:141.
38. ACOG Committee on Practice Bulletins - Obstetrics. Practice Bulletin No. 137: Gestational diabetes mellitus. Obstetric Gynecology 2013;122:406-16.
39. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. Diabetes 1979;28:1039-57.
40. SIGN. 116 Management of diabetes: A national clinical guideline. Scotland: Healthcare Improvement Scotland, 2010 (Updated 2017).
41. Scott DA, Loveman E, McIntyre L, et al. Screening for gestational diabetes: a systematic review and economic evaluation. Health Technology Assessment 2002;6:1-161.
42. Coustan DR, Lowe LP, Metzger BE, et al. The Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study: paving the way for new diagnostic criteria for gestational diabetes mellitus. American Journal of Obstetrics and Gynecology 2010;202:654. e1654. e6.
43. Crowther CA, Hiller JE, Moss JR, et al. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. New England Journal of Medicine 2005;352:24772486.
44. Landon MB, Spong CY, Thom E, et al. A multicenter, randomized trial of treatment for mild gestational diabetes. New England Journal of Medicine 2009;361:1339-1348.
45. US Preventative Services Task Force. Final Recommendation Statement: GDM, Screening. Available at:

https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/gestational-diabetes-mellitus-screening Accessed: 29.10.19, 2014.
46. Association AD. 14. Management of diabetes in pregnancy: standards of medical care in diabetes—2019. Diabetes Care 2019;42:S165-S172.
47. Zhang M, Zhou Y, Zhong J, et al. Current guidelines on the management of gestational diabetes mellitus: a content analysis and appraisal. BMC pregnancy and childbirth 2019;19:200.
48. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine 2011;155:52936.
49. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Annals of Internal Medicine 2019;170:5158.
50. Higgins J, Sterne J, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. Chandler J, McKenzie J, Boutron I, Welch V (editors). Cochrane Methods. Cochrane Database of Systematic Reviews 2016.
51. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. British Medical Journal 2016;355:i4919.
52. Scottish Intercollegiate Guidelines Network. Search Filters: randomised controlled trials; observational studies. Available at: http://www.sign.ac.uk/search-filters.html. Last accessed: 27/06/2020.
53. Dennedy MC, Avalos G, O'Reilly MW, et al. ATLANTIC-DIP: Raised maternal body mass index (BMI) adversely affects maternal and fetal outcomes in glucose-tolerant women according to International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria. Journal of Clinical Endocrinology and Metabolism 2012;97:E608E612.
54. Thaware PK, McKenna S, Patterson CC, et al. Untreated Mild Hyperglycemia During Pregnancy and Anthropometric Measures of Obesity in Offspring at Age 5-7 Years. Diabetes Care 2015;38:1701-6.
55. Farrar D, Fairley L, Santorelli G, et al. Association between hyperglycaemia and adverse perinatal outcomes in south Asian and white British women: analysis of data from the Born in Bradford cohort. The Lancet Diabetes & Endocrinology 2015;3:795-804.
56. Farrar D, Simmonds M, Bryant M, et al. Hyperglycaemia and risk of adverse perinatal outcomes: Systematic review and meta-analysis. BMJ (Online) 2016;354 (no pagination).
57. Lopez del Val T, Alcazar Lazaro V, Garcia Lacalle C, et al. Fasting glucose in the first trimester: An initial approach to diagnosis of gestational diabetes. Endocrinologia, Diabetes y Nutricion 2019;66:11-18.

58. Berntorp K, Anderberg E, Claesson R, et al. The relative importance of maternal body mass index and glucose levels for prediction of large-for-gestational-age births. BMC Pregnancy and Childbirth 2015;15 (1) (no pagination).
59. Biri A, Korucuoglu U, Özcan P, et al. Effect of different degrees of glucose intolerance on maternal and perinatal outcomes. The Journal of Maternal-Fetal & Neonatal Medicine 2009;22:473-478.
60. Cheng YW, Block-Kurbisch I, Caughey AB. Carpenter-Coustan criteria compared with the national diabetes data group thresholds for gestational diabetes mellitus. Obstetrics & Gynecology 2009;114:326-332.
61. Corrado F, Benedetto AD, Cannata ML, et al. A single abnormal value of the glucose tolerance test is related to increased adverse perinatal outcome. The Journal of Maternal-Fetal & Neonatal Medicine 2009;22:597-601.
62. Delibas IB, Tanriverdi S, Cakmak B. Does reactive hypoglycemia during the 100 g oral glucose tolerance test adversely affect perinatal outcomes? Ginekologia polska 2018;89:25-29.
63. Donovan LE, Edwards AL, Savu A, et al. Population-Level Outcomes with a 2-Step Approach for Gestational Diabetes Screening and Diagnosis. Canadian Journal of Diabetes 2017;41:596-602.
64. Jiang S, Chipps D, Cheung WN, et al. Comparison of adverse pregnancy outcomes based on the new IADPSG 2010 gestational diabetes criteria and maternal body mass index. Australian & New Zealand Journal of Obstetrics & Gynaecology 2017;57:533-539.
65. Berggren EK, Mele L, Landon MB, et al. Perinatal outcomes in Hispanic and nonHispanic white women with mild gestational diabetes. Obstetrics & Gynecology 2012;120:1099104.
66. Meek CL, Lewis HB, Patient C, et al. Diagnosis of gestational diabetes mellitus: falling through the net. Diabetologia 2015;58:2003-2012.
67. Miyakoshi K, Tanaka M, Saisho Y, et al. Pancreatic beta-cell function and fetal growth in gestational impaired glucose tolerance. Acta Obstetricia et Gynecologica Scandinavica 2010;89:769-775.
68. Beksac MS, Tanacan A, Hakli DA, et al. Use of the 50-g glucose challenge test to predict excess delivery weight. International Journal of Gynaecology & Obstetrics 2018;142:6165.
69. Berggren EK, Boggess KA, Stuebe AM, et al. National Diabetes Data Group vs Carpenter-Coustan criteria to diagnose gestational diabetes. American journal of obstetrics and gynecology 2011;205:253. e1-253. e7.
70. Davis EM, Scifres CM, Abebe K, et al. Comparison of Birth Outcomes by Gestational Diabetes Screening Criteria. American journal of perinatology reports 2018;8:e280-e288.
71. Ezell JM, Peters RM, Shill JE, et al. Association between Prenatal One-Hour Glucose Challenge Test Values and Delivery Mode in Nondiabetic, Pregnant Black Women. Journal of pregnancy 2015;2015:835613.

72. Noor N, Ferguson KK, Meeker JD, et al. Pregnancy phthalate metabolite concentrations and infant birth weight by gradations of maternal glucose tolerance. International journal of hygiene and environmental health 2019;222:395-401.
73. Verd S, de Sotto D, Fernandez C, et al. The Effects of Mild Gestational Hyperglycemia on Exclusive Breastfeeding Cessation. Nutrients 2016;8:19.
74. Pawelec M, Karmowski A, Krzemieniewska J, et al. The clinical and financial effects of replacing the 1 h 50 g screening test for gestational diabetes mellitus by the stick method. Advances in Clinical and Experimental Medicine 2009;18:601-607.
75. Kosus A, Kosus N, Turhan NO. Gestational diabetes: comparision of the carpenter and the coustan thresholds with the new thresholds of Turkish women and implications of variations in diagnostic criteria. Journal of Maternal-Fetal & Neonatal Medicine 2012;25:616-22.
76. Maesa JM, Fernandez-Riejos P, Sanchez-Margalet V, et al. Fasting Glycemia as Screening Tool to Rule-Out Gestational Diabetes in Low-Risk Population. Clinical Laboratory 2018;64:461-465.
77. Ohara R, Obata-Yasuoka M, Abe K, et al. Effect of hyperemesis gravidarum on gestational diabetes mellitus screening. International Journal of Gynaecology & Obstetrics 2016;132:156-8.
78. Temming LA, Tuuli MG, Stout MJ, et al. Diagnostic ability of elevated 1-h glucose challenge test. Journal of Perinatology 2016;36:342-6.
79. Saeedi M, Hanson U, Simmons D, et al. Characteristics of different risk factors and fasting plasma glucose for identifying GDM when using IADPSG criteria: a crosssectional study. BMC Pregnancy & Childbirth 2018;18:225.
80. Benhalima K, Van Crombrugge P, Moyson C, et al. A Modified Two-Step Screening Strategy for Gestational Diabetes Mellitus Based on the 2013 WHO Criteria by Combining the Glucose Challenge Test and Clinical Risk Factors. Journal of Clinical Medicine 2018;7:13.
81. Ryser Ruetschi J, Jornayvaz FR, Rivest R, et al. Fasting glycaemia to simplify screening for gestational diabetes. BJOG: An International Journal of Obstetrics & Gynaecology 2016;123:2219-2222.
82. Khalafallah A, Phuah E, Al-Barazan AM, et al. Glycosylated haemoglobin for screening and diagnosis of gestational diabetes mellitus. BMJ Open 2016;6:e011059.
83. Gingras V, Rifas-Shiman SL, Switkowski KM, et al. Mid-pregnancy fructosamine measurement-Predictive value for gestational diabetes and association with postpartum glycemic indices. Nutrients 2018;10 (12) (no pagination).

84. Iimura Y, Matsuura M, Yao Z, et al. Lack of predictive power of plasma lipids or lipoproteins for gestational diabetes mellitus in Japanese women. Journal of Diabetes Investigation 2015;6:640-6.
85. Benhalima K, Van Crombrugge P, Moyson C, et al. The Sensitivity and Specificity of the Glucose Challenge Test in a Universal Two-Step Screening Strategy for Gestational Diabetes Mellitus Using the 2013 World Health Organization Criteria. Diabetes Care 2018;41:e111-e112.
86. Benhalima K, Van Crombrugge P, Moyson C, et al. Risk factor screening for gestational diabetes mellitus based on the 2013 WHO criteria. European Journal of Endocrinology 2019;180:353-363.
87. Benhalima K, Van Crombrugge P, Verhaeghe J, et al. The Belgian Diabetes in Pregnancy Study (BEDIP-N), a multi-centric prospective cohort study on screening for diabetes in pregnancy and gestational diabetes: Methodology and design. BMC Pregnancy and Childbirth 2014;14 (1) (no pagination).
88. Theriault S, Forest JC, Masse J, et al. Validation of early risk-prediction models for gestational diabetes based on clinical characteristics. Diabetes Research and Clinical Practice 2014;103:419-425.
89. van Leeuwen M, Opmeer BC, Zweers EJ, et al. Estimating the risk of gestational diabetes mellitus: a clinical prediction model based on patient characteristics and medical history. BJOG: An International Journal of Obstetrics & Gynaecology 2010;117:69-75.
90. Farrar D, Simmonds M, Bryant M, et al. Risk factor screening to identify women requiring oral glucose tolerance testing to diagnose gestational diabetes: A systematic review and meta-analysis and analysis of two pregnancy cohorts. PLoS ONE 2017;12 (4) (no pagination).
91. Brown J, Alwan NA, West J, et al. Lifestyle interventions for the treatment of women with gestational diabetes. Cochrane Database of Systematic Reviews 2017.
92. Brown J, Grzeskowiak L, Williamson K, et al. Insulin for the treatment of women with gestational diabetes. Cochrane Database of Systematic Reviews 2017;2017 (11) (no pagination).
93. Brown J, Martis R, Hughes B, et al. Oral anti-diabetic pharmacological therapies for the treatment of women with gestational diabetes. Cochrane Database of Systematic Reviews 2017;2017 (1) (no pagination).
94. Palatnik A, Mele L, Landon MB, et al. Timing of treatment initiation for mild gestational diabetes mellitus and perinatal outcomes. American journal of obstetrics and gynecology 2015;213:560.e1-8.
95. Hernandez TL, Van Pelt RE, Anderson MA, et al. Women With Gestational Diabetes Mellitus Randomized to a Higher-Complex Carbohydrate/Low-Fat Diet Manifest Lower Adipose Tissue Insulin Resistance, Inflammation, Glucose, and Free Fatty Acids: a Pilot Study. Diabetes care 2016;39:39-42.

96. Reynolds RM, Denison FC, Juszczak E, et al. Glibenclamide and metfoRmin versus stAndard care in gEstational diabeteS (GRACES): a feasibility open label randomised trial. BMC pregnancy and childbirth 2017;17.
97. Senat MV, Affres H, Letourneau A, et al. Effect of Glyburide vs subcutaneous insulin on perinatal complications among women with gestational diabetes a randomized clinical trial. JAMA - Journal of the American Medical Association 2018;319:1773-1780.
98. Sklempe Kokic I, Ivanisevic M, Biolo G, et al. Combination of a structured aerobic and resistance exercise improves glycaemic control in pregnant women diagnosed with gestational diabetes mellitus. A randomised controlled trial. Women and Birth 2018;31:e232-e238.
99. Rowan JA, Rush EC, Plank LD, et al. Metformin in Gestational Diabetes: the Offspring Follow-up (MiG TOFU): body Composition and Metabolic Outcomes at 7-9 Years of Age. Obstetrical & gynecological survey 2018;73:565-567.
100. Pellonpera O, Ronnemaa T, Ekblad U, et al. The effects of metformin treatment of gestational diabetes on maternal weight and glucose tolerance postpartum - A prospective follow-up study. Acta Obstetricia et Gynecologica Scandinavica 2016;95:7987.
101. Hernandez TL, Van Pelt RE, Anderson MA, et al. A higher-complex carbohydrate diet in gestational diabetes mellitus achieves glucose targets and lowers postprandial lipids: a randomized crossover study. Diabetes care 2014;37:1254-1262.
102. Casey BM, Mele L, Landon MB, et al. Does maternal body mass index influence treatment effect in women with mild gestational diabetes? American journal of perinatology 2015;32:93-100.
103. Huhtala, Tertti K, Pellonpera O, et al. Amino acid profile in women with gestational diabetes mellitus treated with metformin or insulin. Diabetes research and clinical practice 2018;146.
104. Trout KK, Homko CJ, Wetzel-Effinger L, et al. Macronutrient composition or social determinants? Impact on infant outcomes with gestational diabetes mellitus. Diabetes Spectrum 2016;29:71-78.
105. Barrett HL, Gatford KL, Houda CM, et al. Maternal and neonatal circulating markers of metabolic and cardiovascular risk in the metformin in gestational diabetes (MiG) trial: responses to maternal metformin versus insulin treatment. Diabetes care 2013;36:529536.
106. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898.
107. Frøslie KF, Røislien J, Qvigstad E, et al. Shape information from glucose curves:

functional data analysis compared with traditional summary measures. BMC medical research methodology 2013;13:6-6.