UK National Screening Committee

Screening for congenital adrenal hyperplasia

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

Version: 2.0

Glen Wilson, Acting Consultant in Public Health

June 2015

The UK NSC advises Ministers and the NHS in all four UK countries about all aspects of screening policy. Its policies are reviewed on a 3 yearly cycle. Current policies can be found in the policy database at http://www.screening.nhs.uk/policies and the policy review process is described in detail at http://www.screening.nhs.uk/policyreview

Template v1.2, June 2010

Plain English Summary

Congenital adrenal hyperplasia (CAH) is caused by altered genes from both mother and the father. It can cause serious illness in babies' shortly after birth and as they get older. A child with CAH may lack the steroid hormone (a type of chemical produced naturally in the body) cortisol that manages the amount of water and salt in the body. Boys and girls with CAH may also have too much of a particular hormone (androgen) that can increase the development of male characteristics.

CAH can affect each infant differently. The symptoms are likely to present in one of the three following ways:

- Classic salt-wasting. This is where the body loses salt and it can cause dehydration and low blood pressure. In severe cases, some organs may stop working properly, and coma or even death may occur. In some female infants, androgens may be present in such quantities that their genitalia may look like a boy's.
- Classic simple-virilising. This is a milder form of CAH where the salt balance is generally better but gender identification may still be difficult.
- Non-classic. Symptoms aren't usually present at birth for this form of the condition. The child may experience abnormal sexual development in later life such as early puberty and inability to produce children.

Newborn screening has been suggested as it is said it might help prevent serious illness in those babies with the severe salt wasting types of the condition. It may also benefit those with the milder forms. The current UKSNC recommendation is that screening should not be recommended. This update review focusses on some of the areas of concern found in the last review.

The conclusions are:

- Study data suggests around 40 babies are born each year with CAH in the UK. Some evidence suggests CAH is more common in people of an Asian background. More studies are needed to better understand whether this is true.

- Studies show the current screening test (using 17-OHP immunoassay) incorrectly identifies a large number of babies as having CAH.

- Studies also show the accuracy of this test was much poorer in babies born early and newborn babies with a low birth weight. This means that affected babies may be missed through screening.

- There was evidence that screening might not reduce the deaths related to CAH and takes place too late benefit people with some types of CAH.

The evidence found in the update would suggest that the recommendation to not screen should not be changed.

Executive Summary

Congenital adrenal hyperplasia (CAH) is an inherited autosomal recessive disorder that affects the enzymes that enable conversion of cholesterol to cortisol at the adrenal glands. The consequence of this is an excess of androgens and a deficit of mineralocorticoids.

CAH can be caused by one of a number of single gene mutations that lead to deficiencies of the 21-hydrolyase enzyme a or, much less commonly, 11β-hydroxylase, 3 β-hydroxysteroid dehydrogenase, 17α-hydroxylase and Steroid Acute Regulatory protein.

The penetrance of the disease is associated with the causal mutation and degree of enzyme disruption. CAH cases are roughly described as one of three phenotypes. The two classic forms of the disease (salt-wasting and simple –virilising) are symptomatic at birth (or shortly after). The more severe salt-wasting form includes infants at risk of adrenal crisis which can lead coma and death, caused by a significant salt imbalance. In contrast, the simple-virilising form's salt balance related symptoms are less profound. In both cases, androgen excess can cause a range of sexual development issues, ambiguous genitalia in females will often be the most immediate symptom- others can include very early puberty or precocious puberty in later life. The non-classic form of the condition is described as a case where no symptoms present in the newborn, these cases may go on to experience symptoms related to sexual development in later life.

The current UKNSC recommendation is to not offer screening for CAH as part of the current newborn bloodspot screening programme. This recommendation is based on a systematic review (Khalid et al., 2010) done on behalf of the UKNSC. The review's main conclusion was on the uncertainty on the screening test; additional concerns were raised about the UK prevalence and acceptability of the screening test.

The 2015 UKNSC update is a rapid review of the 3 clinical questions highlighted in the original review. The review concludes that the volume, quality and consistency of the evidence published does not challenge the conclusions made in the 2010 review.

The findings for each of the key questions are as follows:

What is the incidence of congenital adrenal hyperplasia in the UK population, including discrete subgroups of the population?

- One British Paediatric Surveillance Unit (BPSU) study has been published since the last review. Although the study does suggest the prevalence is within the range seen internationally, there are limitations around the short follow up and the limited depth in which the study could explore incidence within discrete subgroups of the UK population

What evidence exists on the accuracy of the 17-OHP immunoassay?

- The PPV of the test remains a concern, with most studies reporting a false positive which was elevated enough to cause a resultant low PPV. Although no studies undertaken in the UK were identified, international studies consistently show a PPV of 1.5-2.5%. While there is no evidence to suggest that this low PPV would also be the case in the UK, the number of studies reporting similar low outcomes (some of which have analogous incidence rate as in the UK) and the consistently reported elevated false positive rate, would suggest that it is likely to the be similar. Additionally, the test performance was significantly lower in preterm and neonates with a low birth weight.

- Despite popular opinion that the screening test is (or close to) 100% sensitive, a number of studies reported concerns about missed cases and found that the sensitivity is also lower in the non-salt wasting phenotypes.
- Tandem mass spectrometry and repeat bloodspot tests >8days have shown to improve the PPV but the evidence to suggest either is a viable option in a screening programme is limited. Further research on these and other newis required in a large unselected population

Is the 17-OHP immunoassay a suitable population screening test?

- The time cut-offs used in the test appears to mirror that currently used within the dried blood spot programme.
- There are no data on the acceptability of the diagnostic and therapeutic odyssey in a CAH screening programme.
- A French study which retrospectively considered the test accuracy and implications of incorrect results noted that a false negative outcome could have negative implications on a clinical diagnosis.

Introduction

This review will assess the volume, quality and consistency of the evidence for newborn screening for congenital adrenal hyperplasia (CAH), the term given to a number of autosomal recessive disorders of steroid metabolism, published since the last UKNSC review in 2010.

CAH is characterised by impaired adrenal cortisol biosynthesis and an associated androgen excess. (Sharma et al., 2014) CAH can be caused by deficiencies of any of the five enzymes involved in cortisol synthesis: 21-hydroxylase, 11β-hydroxylase, 3 β-hydroxysteroid dehydrogenase, 17α-hydroxylase and Steroid Acute Regulatory protein, with over 90% of cases resulting from 21-hydroxylase deficiency. (Huynh et al., 2009)

Mutations in the genes coding for these enzymes can cause a substantial reduction in enzyme activity, and an associated decrease in cortisol levels together with an increase in levels of adrenocorticotrophic hormone (ACTH). High levels of ACTH, normally inhibited through a negative feedback loop as cortisol levels rise, results in hyperplasia of the adrenal cortex. (Huynh et al., 2009)

There are three broad phenotypes of CAH:

- Classic salt-wasting
- Classic simple-virilising
- Non-classic (late-onset).

They differ in terms of hormone levels, clinical features and age at presentation. (Huynh et al., 2009) Clinically, the salt-wasting phenotype is defined as a sodium concentration of less than 125 mmol/L, the simple virilising form is defined as prenatal virilisation of external genitalia in girls or symptoms before 5 years of age in both sexes, and the non-classic form is defined as onset of symptoms after 5 years of age. All three phenotypes can also be determined genetically. (Gidlof et al., 2013)

In 2010, a detailed evidence review of the suitability of CAH for inclusion in newborn screening was undertaken on behalf of the UK National Screening Committee (NSC). (Khalid et al., 2010). The review assessed the available evidence up to 2008 against 18 criteria to decide if CAH should be included in newborn bloodspot screening programme:

- The condition should be an important public health problem (met)
- The epidemiology and natural history of the condition should be known (not met)
- All effective primary prevention interventions should have been implemented (met)
- There should be a simple, safe, precise and validated screening test (not met)
- The distribution of test values in the population should be known (met)
- The test should be acceptable to the population (partially met)
- If the test is for mutations, criteria should be clearly set out (not applicable)
- There should be agreed policy on diagnostic testing of screen positive individuals (met)
- There should be an effective treatment or intervention (met)
- There should be evidence based treatment policies (met)

- Clinical management should be optimised by all health care providers (not met)
- There should be evidence from high quality randomised controlled trials (not met)
- The screening programme should be clinically, socially and ethically acceptable to health professionals and the public (partially met)
- The benefit of the screening programme should outweigh the harm (met)
- The opportunity cost of the screening programme should be economically balanced (not met)
- Monitoring and management of the screening programme should be to an agreed set of quality assurance standards (partially met)
- Adequate staffing and facilities should be available (partially met)
- All other options for managing the condition should have been considered (met)

The 2010 review highlighted that the key gaps in the evidence to support neonatal screening for CAH related to burden of disease and test performance. The latter, in particular, impacts on a number of the screening criteria, with concerns relating to test specificity and positive predictive value likely to impact on the acceptability of the screening programme as a whole. In addition, cost-effectiveness was not demonstrated in the previous review. Newborn screening for CAH was not recommended by the NSC following the 2010 review.

While the lack of clinical guidelines and variation in management across the UK was recognised in the 2010 review, the development of clinical guidance was not viewed as a major barrier. Similarly, it was considered that quality standards, systems for monitoring and managing the programme and more formal networks could be developed.

The current review will consider whether the volume and direction of the evidence produced since the 2010 review suggests that the conclusions made in that report should be reconsidered. From the 2010 review, four main criteria have been selected as the key conclusions that the recommendation to not screen was based on. Key questions have been developed (see table below) that address each of these conclusions, with particular focus given to fundamental areas the 2010 review identified as uncertain, or supported by insufficient evidence.

Criterion	**Key Questions (KQ)**
2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, disease marker, latent period or early symptomatic stage	What is the incidence of congenital adrenal hyperplasia in the UK population, including discrete subgroups of the population?
5. There should be a simple, safe, precise and validated screening test	What evidence exists on the accuracy of the 17-OHP immunoassay?
6. The test should be acceptable to the population	
14. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is clinically, socially and ethically acceptable to health professionals and the public.	Is the 17-OHP immunoassay a suitable population screening test?

It was agreed that the key questions would need to be adequately answered in order to complete a cost-effectiveness review. Therefore, a health economic assessment will only be undertaken if new evidence is identified that answers the key questions outlined above.

Appraisal against UK NSC Criteria

These criteria are available online at http://www.screening.nhs.uk/criteria.

2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, disease marker, latent period or early symptomatic stage

The previous evidence review for CAH screening noted a lack of robust evidence on CAH prevalence in the UK, with estimates ranging from 1 in 6,200 to 1 in 25,000 live births. (Khalid et al., 2010)

Description of evidence

The evidence search for the update review identified a UK study and 13 international studies on the incidence, distribution and clinical presentation of CAH, published since the 2010 review (summarised in table 1). 3 studies were done in an unselected population; the others were taken from screened infants. The range from studies of CAH identified solely from screening being 1:9030 in UAE to 1:25,550 in Estonia. The inter-quartile range for the studies of CAH screening was between 1:15,931 to 1:19,939.

In addition to the studies outlined in Table 1, a recent publication cited a birth prevalence rate in the United States of America for CAH of 1:20,800, with a range of 1:15,900 to 1:27,000 across different states, though limited information was provided on distribution and clinical features, and the data were from 1991 to 2000, at which point CAH screening was not carried out in the majority of states. (Hertzberg et al., 2011)

Table 1. Summary of CAH incidence, distribution and clinical presentation across screened & unscreened populations

Country	Day of screening	Median age at clinical presentation	% of cases symptomatic at diagnosis	Number of CAH cases detected & time period	CAH incidence (no. screened)	Comments
Sweden (Gidlof et al., 2013)	Unreported	-	Of 274 cases screened*, 84% detected through screening, 16% detected clinically at a later date (false negatives).**	606 in total (including those before introduction of screening) 292* screened 92 SW female 73 SW male 71 NSW female 56 NSW male 1910-2011	1:8,800 (2,737,932 screened)	* narrative states 274 screened; data presented in Figure 2 indicates 292 screened. Subsequent publication confirms 274 (Gidlof, 2014) **24/43 (56%) of false negatives had non classic CAH The overall female-male ratio of individuals detected via screening was close to 1 (115 vs 116) In the whole cohort it was 1.25, which the authors attribute to higher rates of detection of mild non-salt wasting forms in later life in females not higher rates of early mortality from SW forms in males.
France (Coulm et al., 2012)	3 or 4	7 days	28.5% of cases had been detected clinically before screening results were available and 19.3% were detected due to family history. Screening was useful for detecting 42.3% of total cases. 6.5% of cases were missed by screening; no information were available on the remaining 3.4%.	383 358 had positive screening results 25 had negative screening results, detected later 1996-2003	1:15,699 (6,012,798 screened)	The majority of infants for whom screening was useful (correct positive test in infants with no symptoms) in diagnosis were males with the salt wasting form. An additional 12 infants had nonclassic CAH not included in totals
Czech Republic	3-4	-	16.6% clinical suspicion of CAH prior to screening result	48	1:11,354 (545,026	

(Votava et al., 2012)				28 male 19 female 2006-2010	screened)	
USA (Chan et al., 2013)	3 (first screen) 8-14 (second screen)	-	-	29 (first screen) Plus additional 17 (second screen)	1:24,766 [first screen] 1:17,789 [first and second screen combined] (693,751 screened)	This paper reports experience from Colorado, USA, where a second newborn screening test is in place for CAH
Brazil [Santa Catarine] (Nascimento et al., 2014)	7.3 (mean)	-	-	50 21 SW female 16 SW male 13 SV 2001-2010	1:14,967	
Brazil [Minas Gerais] (Pezzuti et al., 2014)	5-9	-	50% (four females clinically suspected to have CAH before screening test). Further female had symptoms but no clinical suspicion.	8 3 SW female 3 SW male 2 SV female 2007-2008	1:19,927	Same population reported by Barra, 2012
Brazil [Minas Gerais] (Barra et al., 2012)	6 (median)	-	-	8 5 female 3 male 75% SW 2007-2008	1:19,939[£]	Same population reported by Pezzuti, 2014
Brazil [Sao Paulo] (Hayashi et al., 2011)	2-13	-	-	7 5 SW male 1 SW female 1SV female	1:10,460	Retrospective analysis of CAH screening pilot

				2009-2010		
Cuba (Gonzalez et al., 2013)	6	-	-	39 15 SV 16 SW 8 NC 2005-2010	1:15,931 (621,303 screened)	
United Arab Emirates (Al Hosani et al., 2014)	3	-	-	42 2007-2011	1:9,030 (750,365 screened)	Authors note incidence may be increased by high rates of consanguineous marriage in this population
Japan (Morikawa et al., 2014)	4-6	-	11 females symptomatic before screening test*	26* 19 female 7 male 20 SW 4 SV 2 NC 1982-2010	1:19,160 or 1:20,756 [classical CAH] (498,147 screened)	*Discrepancy in total number of cases presented and narrative – varies from 25 to 26.
Unscreened populations						
Great Britain (Khalid et al., 2012)	N/A	Female 0 days Male 14 days Overall 1 day	Females: 86% (77% virilisation, 7% salt wasting or adrenal crisis, 2% adrenal insufficiency). 14% presented due to affected sibling. Males: 85% (73% salt wasting or adrenal crisis, 6% adrenal insufficiency, 6% incomplete masculinisation). 15% presented due to affected sibling.	144 58 male 77 female 9 from August 2007 excluded from age stratification by study authors 2007-2009	1:18,000 (not reported)	See key question below for more detailed appraisal
Croatia (Dumic et al., 2009)	N/A	SW female 8 days SW male 3.8 weeks	All live born cases symptomatic at diagnosis. 100% females had ambiguous	37 12 SW female 8 SW male 7 SV female	1:14,403 (532,942 births)	Prenatal diagnosis[$] resulted in elective abortion of three female fetuses, included in numbers presented here

		SV female 7.2 months SV male 4.8 years	genitalia; 66% of SW females additionally had adrenal crises 100% SW boys had adrenal crises; 100% SV boys had accelerated growth & development	7 SV male 3 aborted female fetuses 1995-2006		Non-classic CAH not reported
Estonia (Liivak et al., 2008)	N/A	SW males 30 days SW females 2 days	100%	20 SW males 7 SW females 7 SV males 1 SV females 6	1:25,500 (not reported)	Uncertainty in incidence; authors suggest the incidence since regaining independence (1:16,100) may be more accurate

[$] Prenatal diagnosis was performed by karyotyping and analysis of the CYP21 gene on a sample obtained by chorionic villous sampling at 10–12 weeks' gestation.

[£] Difference in incidence from Pezzuti 2014 is because of rounding in study

SW = salt wasting; SV = simple virilising; NSW = non salt wasting

Key question: What is the incidence of congenital adrenal hyperplasia in the UK population, including discrete subgroups of the population?

A study undertaken through the British Paediatric Surveillance Unit (BPSU), published in Khalid et al., 2012 and Knowles et al., 2014 specifically reported on the incidence of CAH in Great Britain. The prospective surveillance study ran over a 24 month period and identified a total of 144 infants born with CAH, 86 (60%) who presented clinically in the first year of life and 58 (40%) over one year. Based on national birth rates, the authors estimated a CAH incidence of approximately 1:18000.

The frequency of CAH subtypes also mirrored that reported internationally, with over 90% of cases due to 21-hydroxylase deficiency and the majority of the remainder due to 11β-hydroxylase deficiency.

There was a slight female excess of cases, though unlike international experiences, this was not a statistically significant difference. There did appear to be a strong association with ethnicity, with 28% of CAH cases being of Asian ethnicity, despite 2013 ONS statistics suggesting births to those of Asian ethnicity comprising a markedly lower proportion of the general population of Great Britain.

A Kaplan-Meier plot was used to illustrate the age at which CAH was diagnosed in the 144 cases, split by sex. The plot illustrated that girls were generally diagnosed earlier than boys, and most often on the first day of life. Age at diagnosis was presented in the context of current UK standards for the reporting of newborn screening results, which should be available by day 14 after birth. A total of 6% of girls and 50% of boys remained undiagnosed at day 14, indicating that these infants may have benefited from newborn screening. Of particular note is that of the 27 newborns that presented with salt-wasting crises, widely viewed as the most severe and life-threatening clinical feature of CAH, 18 (66%) presented on or after day 14, when screening results would have been available.

This is the only study which directly estimates incidence in Great Britain (note that the study did not cover Northern Ireland therefore no estimate is available for this population of the UK). The existence of only one study, conducted over a relatively short period with a short-term follow up period, represents limited evidence on UK incidence. The study was however conducted using sound methodology, based on a well-established national surveillance system. The external validity of the study is supported by the consistency of the reported incidence within the range of rates estimated globally, at between 1:10,000 to 1:20,000, (Speiser et al., 2010) and the range of studies reported in table 1.0. However, higher incidence of CAH in those of Asian ethnicity was not a finding identified in the literature reviewed in the Khalid et al., 2010 study or in the recent international studies outlined in Table 1.

Although the primary rationale for newborn screening for CAH is to enable early recognition and treatment to prevent mortality and morbidity from salt wasting crises, a recent study from the North West of England found no evidence that undiagnosed CAH had contributed to infant mortality over a 12 year period. (Hird et al., 2013)

Another study from Britain provides evidence of potential impact on morbidity and quality of life. (Knowles et al., 2013) A prospective paediatric surveillance study which ran from 2007-2009 identified 58 children who presented with CAH after the age of one year in England, Scotland and Wales, equivalent to 0.23 per 100,000 or around 30 cases per year. There was little evidence of difference in rates of presentation by sex, with approximately equal numbers of both genders presenting aged one or older during the study period. The median age at

presentation was 5.9 years, and similar to the previously described British study (Khalid et al., 2012) there was an excess of cases among those of Asian ethnicity. (Knowles et al., 2013)

Summary Criterion 2: Partially met. The study of CAH incidence in Great Britain provides some evidence that UK incidence is likely to be within typically quoted international estimates. The evidence was limited by the short-term duration of the study and by limited evidence regarding incidence within discrete subgroups of the UK population.

5. There should be a simple, safe, precise and validated screening test

The primary biomarker for CAH screening is 17-hydroxyprogesterone (17-OHP). 21-hydrolyase is the enzyme which converts 17-OHP into the cortisol substrate 11-deoxycortisol. 21-hydroxylase deficiency, responsible for the vast majority of cases of CAH, leads to failure of this step of the pathway and an accumulation of 17-OHP. The 17-OHP immunoassay is also effective in detecting some but not all cases of 11-hydroxylase deficiency. (Janzen et al., 2012) The remaining three enzymes act at points higher in the pathway, preventing the synthesis of 17-OHP, and therefore the assay is unable to detect CAH caused by deficiency in 3 β-hydroxysteroid dehydrogenase, 17α-hydroxylase or Steroid Acute Regulatory protein, which collectively account for less than 5% of all CAH cases

Various forms of immunoassay have been utilised as primary screening tests for CAH. Early screening tests utilised radioimmunoassays, and more recently enzyme-linked immunosorbent assay (ELISA), though both now have largely been replaced in newborn screening programmes by automated time-resolved dissociation-enhanced lanthanide fluoroimmunoassay (DELFIA) (Speiser et al., 2010). A comparison of CAH detection rates in US states using different screening technologies from 1991-2000 found that ELISA was the only technology associated with a significantly higher CAH detection rate than radioimmunoassay, (Hertzberg et al, 2011) however the number of states screening for CAH and particularly the number utilising ELISA was very low, which led the authors to note an uncertainty in this finding.

The previous evidence review highlighted concerns regarding positive predictive value (PPV) and the false positive rate of the 17-OHP immunoassay. CAH is a relatively rare condition and, as such, relatively small changes in the false positive rate will have a significant effect on the positive predictive value, despite both the specificity and sensitivity of the test appearing to be high. Furthermore, the false positive rate was consistently reported to be higher in premature

and low birth weight infants. (Khalid, 2010)

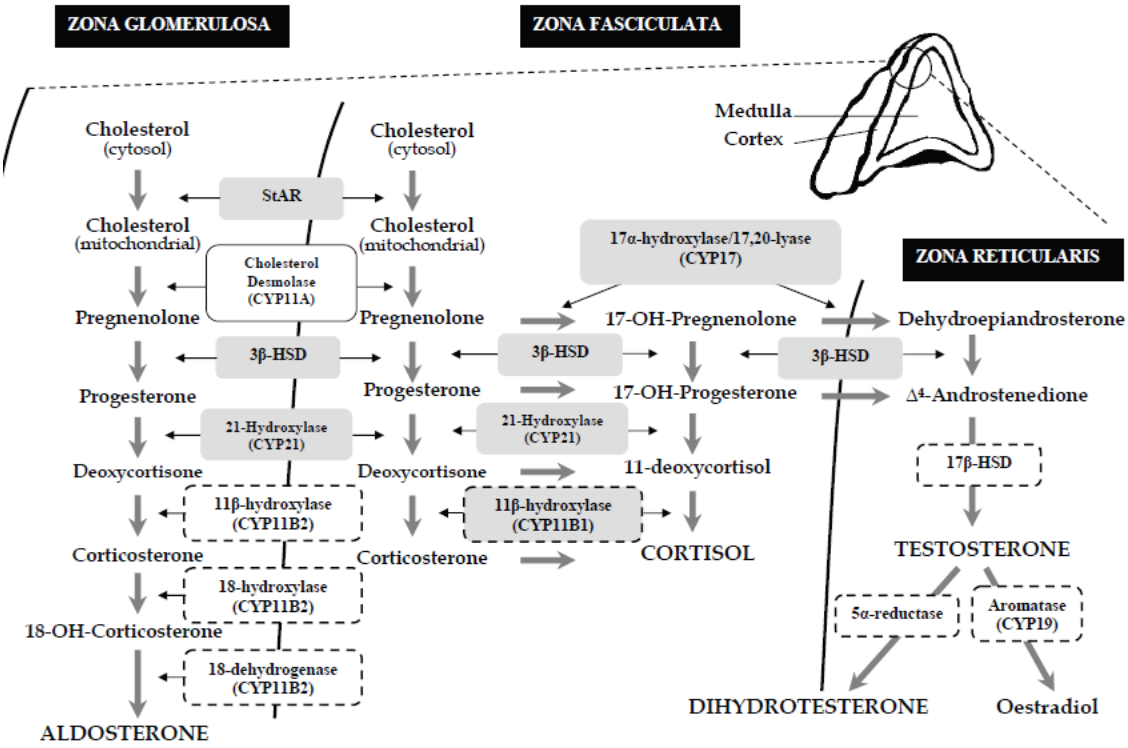

Figure 1 Schematic representation of steroidogenesis.
Shaded boxes highlight the five enzymes implicated in CAH. *Source: Huynh et al., 2009*

Current UKNSC key question

The current review focuses on the evidence relating to the clinical value of newborn screening immunoassays for CAH, prioritising prospective studies of large, unselected or representative populations. In the absence of this type of evidence on the clinical validity, case-control studies assessing the analytical validity of testing strategies have been considered.

Description of the evidence

10 studies were included in the review (see table 2). 5 were cohort studies from which outcomes were reported from pilot screening programmes in Brazil and USA. 3 were retrospective cohort studies reporting outcomes from national screening programmes in France, Cuba and Czech Republic. 1 study was a Swedish longitudinal, prospective study and the final study was a case-series which reported outcomes on missed cases in the German screening programme. The study populations ranged from 67,640 to 6,012,798. No studies were done in a UK population.

The studies reported PPV ranges from 0.4% to 13.4%, with the majority of studies reporting PPVs no higher than 2.3%. The reported false positive rates ranged from 0.1% to 1.24%. Many of the studies noted that the predictive value of the test is associated with the gestational age at delivery and/or birth weight.

The test sensitivity ranged from 71.8% to 100%. The lowest sensitivity was reported in a retrospective cohort study, published as a letter, which covered a 12 year period of newborn screening for CAH in Minnesota, during which time 52 patents with classic CAH were identified and 15 missed by screening, giving a false negative rate of 22.4%. Of the missed cases, 6 were males and 9 were females, with both salt wasting (5 cases) and simple virilising (10 cases) forms of CAH missed. (Sarafoglou et al., 2012a) A similarly high level of false negative newborn screens was reported in Colorado. (Chan et al., 2013)

One study (Gildof et al., 2014) reported differences in the sensitivity of the screening tests done in the Swedish national programme for the detection of the different phenotypes. The sensitivity of the test was highest for the salt wasting form (100%) and lowest in the non-classic phenotype (32.4%), a number of the other studies also noted that the sensitivity was highest for the salt wasting form of the disease.

In a recent French study a significant difference in the performance in preterm and term neonates was reported, with a 1.5% difference in the false positive rate between the two. (Coulm et al., 2012) Although the overall PPV in this study was similar to other studies, at 2.3%, the authors demonstrated that the majority of false positives were in preterm neonates, which meant that PPV increased to 30.1% in an sub-group analysis of only neonates born at term compared to a PPV of just 0.4% in preterm neonates. The study cautioned against using adjusted cut-offs based on gestational age as there was marked overlap in 17-OHP levels between term and preterm neonates, highlighting potential loss of sensitivity. A longitudinal, prospective study from Sweden reported similar findings relating to PPV in term compared to preterm infants. (Gidlof et al., 2014) The overall PPV in that study was relatively high, at 13.4%, but shown to increase markedly to 25.1% for term infants compared to only 1.4% in preterm infants. (Gidlof et al., 2014). In both studies the false positive rate was higher in the pre-term group which would indicate the difference of PPV can be attributed to the test performance and not just a difference in prevalence in pre-term and term infants.

Several studies corroborated earlier research (Khalid et al., 2010) which indicated that use of adjustments for 17-OHP concentration based on prematurity and/birthweight reduced the false positive rate and improved PPV. (Barra et al., 2012; Hayashi et al., 2011; and Gonzalez et al., 2013)

Discussion

The studies included were undertaken in large unselected populations; however none compared the screening test to clinical diagnosis following pathways similar to those done in the UK.

The positive predictive value of the screening test was widely reported to be low however, as positive predictive value is influenced by the prevalence of the condition in a given population, this reduces the applicability of these results to a UK screening programme.

The false positive rate in the reported studies was relatively low but, because of the low incidence rate of CAH, it was acknowledged to be an important factor in the low reported PPVs. The underlying cause of the false positives is thought to be attributable to cross reacting metabolites which are significantly altered in premature infants and by increased 17-OHP in unwell (those with poor kidney liver function) or stressed babies. It's because of the latter reason that the test is not done before 72hours, in most cases. (Votava et al., 2012). As noted in many studies, there is a significant difference in the performance of the screening test in preterm and term infants. Coulm et al., 2012 recommended that newborn screening in preterm

neonates be discontinued in the French screening programme and added caution against using adjusted cut-offs based on gestational age.

A popular perception is that the screening test's sensitivity is 100%, however several studies reported false negative results in international newborn screening pilots an programmes. (Sarafoglou et al., 2012a; Chan et al., 2013; Gidlof et al., 2014 and Schreiner et al., 2008) It was also found that a false negative screening result could lead to delay in diagnosis, even in the presence of clinical features consistent with CAH.

Schreiner et al., 2008 provides further evidence that a negative CAH screening result does not exclude the possibility of CAH. There were five false negatives in the Shreiner et al., 2008 study, two males and three females. Both males and one female were affected by simple virilising CAH and the other two females by salt wasting CAH. As reported in other studies, a false negative screening result delayed diagnosis, even in the presence of clinical features, with clinicians referring back to the screening result in determining not to initiate diagnostics in response to virilising features.

No studies suggested that the screening test could not distinguish between the phenotypes, nor could they predict severity. There was some data that noted that the sensitivity of the test was significantly lower in the non-salt washing phenotypes, particularly in females. Furthermore, Coulm et al., 2012 reported that many female cases were diagnosed in neonatal paediatric examinations and a small proportion of boys (9 of 153) with salt wasting forms presented with symptoms – both occurred before screening test could be offered.

Due to the short term follow up in a number of other studies, and inherent difficulties associated with long term follow-up, it is likely that other cases of false negative screening results may have gone unreported.

Approaches to improving PPV

It has been demonstrated that the false positive rate can be reduced by as much as 40%, without reducing sensitivity, through employing diethyl ether to extract 17-OHP from dried blood spots. Ether extraction separates 17-OHP from some but not all other steroids, improving specificity of the test, while also completely removing the effect of EDTA contamination, which has been implicated as a cause of false positives. (Fingerhut et al., 2009)

Addition of a repeat test has been proposed as a means of improving the positive predictive value and reducing the false positive rate of newborn screening for CAH. A recent study has advocated a second screening sample to be collected on day 8-14, presenting evidence of an improvement in PPV from 0.4% on the first screen to 6.8% following the second screen. (Chan et al., 2013) While this represents a substantial improvement in PPV, collection of a second bloodspot sample from infants on day 8-14 would present a significant logistical challenge and is unlikely to be feasible.

The use of more specific second tier tests has also been suggested to improve overall accuracy. Use of tandem mass spectrometry as such a test is advocated in a number of studies, (Rossi et al., 2011; Dhillon et al., 2011; Janzen et al., 2011; Seo et al., 2014 and Salter et al., 2015) as well as in a recent clinical guideline. (Speiser et al., 2010) The use of a 3 steroid profile reported in one study was shown to reduce the false positive rate by 93%, (Dhillon et al., 2011) while another study reported use of tandem mass spectrometry as a second tier test eliminated all false positive results. (Seo et al., 2014) By contrast, a study from Minnesota found little impact

of tandem mass spectrometry on the false positive rate, and persistence of a high false negative rate, with one case missed for every three identified. (Sarafoglou et al., 2012b).

Further research is ongoing on ways to improve the predictive value of CAH, including recent exploration of genetic testing following a positive primary test. It was noted, however, the data for each of the strategies to improve the positive predictive value of screening was limited. To date, there are no large cohort studies, undertaken in an unselected population, that adequately demonstrate an improvement that is applicable in a UK screening population.

Table 2. Performance of first tier screening tests for CAH

Country	Day of screening	Test	Cut-off value(s)	Sn (%)	Sp (%)	PPV (%)	NPV (%)	Comments
Brazil (Barra et al., 2012)	6 (median)	ELISA Incidence is 1:19,939 159,415 children screened (8 children with CAH; 5 females and 3 males)	<80 nmol/L normal 80-160 nmol/L retest >160 nmol/L medical assessment	100	99.6	2.2	-	By adjusting the cut-offs to 110 nmol/L and 220 nmol/L, a 76% decrease in medical referrals was projected *Same population reported by Pezzuti, 2014*
Brazil (Pezzuti et al., 2014)	5-9	ELISA Incidence is 1:19,939 159,415 children screened (8 children with CAH; 5 females and 3 males)	<80 nmol/L normal 80-160 nmol/L retest >160 nmol/L medical assessment	100	99.7	2.1	-	3% of false positives were low birth weight, premature or both *Same population reported by Barra, 2012*
Brazil (Hayashi et al., 2011)	2-13	Fluoroimmunoassay Incidence not reported. 67,640 children screened	≥20 ng/mL repeat sample	>99%	99.3-99.5	1.5-2.0	99.3-99.5	Ranges refer to adjustments in cut-off values based on birthweight
Cuba (Gonzalez et al., 2013)	6	ELISA Incidence is 1:15,931, 621,303 children screened (39 children with CAH)	>55 nmol/L	-	98.26	0.36	-	

USA (Sarafoglou et al., 2012a)	1-2	Fluoroimmunoassay Incidence not reported. 838,241 children screened (67 children with CAH)	-	77.6*	-	-	-	15/67 cases missed = 77.6% sensitivity
USA (Chan et al., 2013)	3	Fluoroimmunoassay Incidence is 1:17,789. 693,751 children screened (46 children with CAH)	Initially 55 ng/mL for normal birthweight, decreased to 35 ng/mL Higher cut-offs used for low birthweight neonates	71.8	99	0.4	99.9	PPV increased to 3.7% as a result of changing the cut-off level in low birth weight neonates
Germany (Schreiner et al., 2008)	3-4	Not reported	Varied	97.6*	-	-	-	5/214 cases missed = 97.6% sensitivity
France (Coulm et al., 2012)	3	Fluoroimmunoassay or radioimmunoassay Incidence not reported 6,012,798 children screened	≥40 nmol/l (fluoroimmunoassay) or ≥50 nmol/l (radioimmunoassay)	93.5 93.8 term 95.0 preterm	99.7 99.9 term 97.6 preterm	2.3 30.1 term 0.4 preterm	99.9 (all)	Most false positives were preterm and most false negatives were simple virilising form
Czech Republic (Votava 2012)	3-4	Fluoroimmunoassay Incidence is 1:11,354. 545,026 children screened (46 children with CAH)	20 to 200 nmol/L depending on birthweight	98.0	99.5	1.6	-	
Sweden (Gidlof et al., 2014)	2+	Radioimmunoassay (1986-91) Fluoroimmunoassay (1991 onwards). 2,737,932 children screened (231 children with CAH)	Varied over time and different cut-offs used for preterm neonates	84.3 SW 100 SV 79.7 NC 32.4	99.9	13.4 overall 25.1 (term) 1.4 (preterm)	"close to 100%"	

*** Reviewer calculated**

Summary Criterion 5: Not met. The first tier test (immunoassay) has high sensitivity though importantly, is not assured at 100% as widely reported. As a result, some CAH cases are missed through first tier testing, and there is evidence that a negative screening result to can delay clinical investigation, even in the presence of symptoms. Due to short follow up in a number of studies, it is likely that false negatives are underreported and therefore the sensitivity has been overestimated.

Although it was noted that screening programmes are often designed to detect the more severe salt wasting performance, the evidence has shown that the sensitivity for other phenotypes is significantly lower and more so in females than males. Furthermore, one study noted that a number of cases of salt-wasting and simple-virilising cases were detected either through neonatal assessment or symptom based diagnosis before screening could be offered.

Data on PPV and the false positive rate is largely consistent with that reported in the previous review. As CAH is a relatively rare condition the false positive rate (the inverse of the specificity) will have a substantial effect on the PPV. Although most studies report a false positive rate within a range 0.1% to 1.24% the effect on the total number of false positives, with respect to confirmed cases, is very high which means that the absolute number of true positives among those who screen positive will be very low. The FPR can therefore offer a partial explanation of the low positive predictive value reported in the studies, which is independent of variation in the prevalence of the condition and more a reflection of the limitations of the screening test.

More specific evidence has been published in recent years to demonstrate that the false positive rate is increased in premature and/or low birthweight neonates. Both the false positive rate and PPV are expected to be more useful in term neonates, which have led some authors to suggest that screening for CAH is not useful in preterm neonates.

Tandem mass spectrometry as a second tier test has been advocated by a number of studies (and in a recent clinical guideline) as a means of improving the positive predictive value of newborn screening for CAH. The practicalities of use of the technology within the UK screening program require further consideration, as does the evidence from one study which found a persistently high false negative rate, even with a two tier screening protocol.

6. The test should be acceptable to the population

The 17-OHP assay in the studies included in criteria 5 use a similar, if not the same, sampling method as the newborn blood spot screening programme in the UK. However, not all the studies reported a similar timeframe for taking the sample as that currently used in the UK programme (5 to 8 days after delivery). Whether the differences in sample collection is a critical concern for applicability depends on whether there is variation in 17-OHP accumulation over the course of the first week of life among infants with CAH. Authors have speculated that a later time may actually improve predictive utility of the test (Chan et al., 2013). The dried blood spot programme has shown to be acceptable to the population but no evidence was identified that reported outcomes on the acceptability of a variation in time or methods of the current DBS for the purposes of screening for CAH.

Further evidence for this criterion is presented in the preceding section.

Summary Criterion 6: Partially met. The test would not involve additional blood sampling from neonates as the 17-OHP assay is based on the newborn bloodspot, already routinely carried out in the UK to test for other conditions. No evidence was identified that can answer

whether an adaption of the current blood spot programme sampling pathway would be acceptable in order to detect CAH.

Concerns identified in the previous review regarding the low PPV of the first tier and associated harms and costs remain. In addition, evidence is emerging to demonstrate that the sensitivity of the test is not 100%, meaning some affected neonates are missed through first tier testing. A false negative screening result has been shown to delay investigation and diagnosis, even in the presence of symptoms.

A social value judgement will be required to determine whether more good is done through accurately detecting most affected babies, than harm through false positive and false negative screening results. No evidence has been identified that addresses this question in a UK population.

14. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is clinically, socially and ethically acceptable to health professionals and the public

No studies were identified that reported evidence on the opinions and perspectives of patients, patients and health professionals in a CAH screening programme, in the UK. Since the previous review, a clinical guideline on CAH has been produced by The Endocrine Society. (Speiser et al., 2010) The guideline recommends that screening for 21-hydroxylase deficiency should be included in all newborn screening programmes, using a two tier protocol. It further recommends that test cut-offs are stratified by gestational age. Although both recommendations are graded 'strong', the guideline makes clear that both are based on low quality evidence.

An audit of newborn screening practices in Europe was published in 2012, which showed little consensus in the conditions included in national newborn screening panels. (Loeber et al., 2012) This study identified that CAH was the third most-screened for condition in newborn screening panels across Europe, included by 15 countries. European countries screening for CAH comprised Albania, Austria, Belgium, Bosnia-Herzegovina, Bulgaria, Czech Republic, Denmark, France, Germany, Luxembourg, Netherlands, Slovakia, Spain, Sweden and Switzerland.

A surveillance study in the UK noted that the short follow-up period did not enable the paper to conclude whether the identified late clinical presentations of CAH were 'missed' cases of salt wasting and simple virilising phenotypes, or true late onset (non-classic). Knowles et al.,2013 This has important implications for relevance to newborn screening, since the immunoassays used in screening programmes have limited ability to detect non-classic CAH. (Huynh et al., 2009 and Votava et al., 2012)

Summary Criterion 14: No studies were identified that addressed the diagnostic and therapeutic odyssey that parents and patients would experience. Therefore the review cannot make a conclusion on the benefits of an early diagnosis and treatment nor the limitations and harms that would be associated with the expected high number of false positives, if a screening programme was introduced in the UK.

Screening is undertaken in 15 European countries, all 50 American states and several other developed countries and is recommended by The Endocrine Society. Concerns persist regarding test accuracy and one French study was identified which argued for discontinuation of screening for CAH in premature infants due to very low PPV in this group. Several studies

have demonstrated that a proportion of infants are diagnosed due to clinical presentation prior to screening results being available. Most commonly this relates to the simple virilising form in females, but examples of adrenal crises in the salt wasting form have also been reported to occur prior to availability of screening results.

Evidence from the UK is limited, but that which is available indicates that newborn screening for CAH would be unlikely to impact on infant mortality. There is some evidence that late-presenting cases could be identified earlier, though it remains unclear whether screening would accurately identify these cases, particularly if they represent the non-classic CAH phenotype.

Implications for policy

This report assesses newborn screening for CAH against the UK National Screening Committee (UK NSC) criteria for appraising the viability, effectiveness and appropriateness of a screening programme. This topic was last reviewed by the NSC in 2010. (Khalid et al., 2010) The review concluded that the key gaps in the evidence to support neonatal screening for CAH related to burden of disease and test performance.

This review assessed key questions surrounding the UK incidence of CAH, the accuracy of the screening test, the acceptability of the test and the acceptability of the screening programme. Overall, the evidence identified was of low quality, comprising of observational studies. This limited body of evidence provided some additional information to that included in the 2010 review; key findings for each of the criteria are summarised below:

- **Epidemiology of the condition:** A single study estimated the incidence of CAH in the UK was around 1:18,000, which is within typically quoted international estimates. There is little firm evidence regarding incidence within discrete subgroups of the UK population, though incidence appeared disproportionately high in those of Asian ethnicity. This ethnic association has not been reported in other population studies and requires further corroboration.
- **Test accuracy:** The routinely used immunoassay has high sensitivity; however recent studies have reported false negatives, challenging the assertion that test sensitivity is 100%. Furthermore, due to short follow up in a number of studies, it is likely that false negatives are also underreported and therefore sensitivity overestimated. Also, it was noted that the sensitivity of test was lower for the detection of the non-salt washing phenotypes and that presentation of symptoms and detection through physical examination can identify cases before screening is offered.

  The PPV of the test remains a concern, with most studies reporting a false positive rate which was elevated enough to cause a resultant low PPV. Although no studies undertaken in the UK were identified, international studies consistently show a PPV of 1.5-2.5%. While there is no evidence on PPV of the test in the UK, the number of studies reporting similar low outcomes (some of which have analogous incidence rate as in the UK) and the consistently reported elevated false positive rate, would suggest that it is likely to the be similar. A small number of studies have stratified PPV according to gestational age, finding that in term neonates the PPV is significantly higher (at around 25-30%) compared to the PPV in preterm neonates. These findings have prompted calls for screening to be discontinued in preterm neonates.

  Tandem mass spectrometry as a second tier test has been widely advocated and is recommended in a recent clinical guideline as a means of improving the positive predictive value of newborn screening for CAH. The practicalities of use of the technology within the UK screening program require further consideration, as does the evidence from one study which found a persistently high false negative rate, even with a two tier screening protocol.

- **Acceptability of test:** The screening test requires no additional blood sampling from neonates as the 17-OHP assay is based on the newborn bloodspot, already routinely carried out in the UK to test for other conditions.

Fundamental concerns remain regarding the high number of false positive results and resultant low predictive value of a positive test result. The false negative rate – in some studies found to be as high as one in four – means that affected neonates may be missed through screening, and a false negative screening result has been shown to delay investigation and diagnosis, even in the presence of symptoms.

All of the proposed approaches to improving the PPV have associated challenges: adjusting cut-offs for birthweight or gestational age may reduce sensitivity, screening only term infants is not consistent with a universal screening offer and raises ethical considerations and use of second tier tests is costly and time consuming.

A value judgement is required to determine whether more good is done through accurately detecting most affected babies, than harm through false positive and false negative screening results.

- **Acceptability of the screening programme:** Screening is undertaken in 15 European countries, all 50 American states and several other countries and is recommended by The Endocrine Society. Concerns persist regarding test accuracy and one study was identified which argued for discontinuation of screening for CAH in premature infants due to very low PPV in this group. Several studies have demonstrated that a proportion of infants are diagnosed due to clinical presentation prior to screening results being available. Most commonly this relates to the simple virilising form in females, but examples of adrenal crises in the salt wasting form have also been reported to occur prior to availability of screening results.

  Evidence from the UK is limited, but that which is available indicates that newborn screening for CAH would be unlikely to impact on infant mortality. There is some evidence that late-presenting cases could be identified earlier, though it remains unclear whether screening would accurately identify these cases, particularly if they represent the non-classic CAH phenotype.

Implications for research

Given the limited evidence identified for each of the key questions, additional high quality studies in the following areas would be useful in order to resolve uncertainties regarding newborn screening for CAH:

- Studies that provide more a robust estimate of UK incidence, over a longer time period to provide assurance that estimates are representative of true incidence in the UK population and to observe trends over time.
- In particular, it will be important to identify whether the apparent association with ethnicity is a true finding, or due to chance, and if true to explore why this pattern has not been more widely reported.
- Longer follow up in studies conducted within screened populations would be useful to provide assurance that any false negatives would be detected clinically, to provide a more robust estimate of test sensitivity. This should include false positive rates for the various forms of CAH.
- Studies designed to confirm the impact of screening only term neonates on test performance, and particularly on PPV.

Methodology

Search strategy

DATABASE: Medline

SEARCH DATE: 19 January 2015

Publication date limits: January 2008 – 19 January 2015

Prevalence/burden of disease

1 ADRENAL HYPERPLASIA, CONGENITAL/ [Limit to: Publication Year 2008-2015] (1000)

2 STEROID 21-HYDROXYLASE/ [Limit to: Publication Year 2008-2015] (397)

3 "congenital adrenal hyperplasia".ti,.ab [Limit to: Publication Year 2008-2015] (1049)

4 1 or 2 or 3 (1470)

5 EPIDEMIOLOGY/ [Limit to: Publication Year 2008-2015] (1017)

6 incidence.ti,ab [Limit to: Publication Year 2008-2015] (179025)

7 prevalence.ti,ab [Limit to: Publication Year 2008-2015] (181733)

8 4 or 5 or 6 (344050)

9 4 and 8 (155)

****

Test accuracy

10 17-ALPHA-HYDROXYPROGESTERONE/ [Limit to: Publication Year 2008-2015] (276)

11 (17 AND hydroxyprogesterone).ti,.ab [Limit to: Publication Year 2008-2015] (533)

12 10 or 11 (639)

13 (detect$ OR test OR tests OR testing OR screen$).ti,ab [Limit to: Publication Year 2008-2015] (1196420)

14 (sensitiv$ OR specific$).ti,ab [Limit to: Publication Year 2008-2015] (1014869)

15 "predictive value".ti,ab [Limit to: Publication Year 2008-2015 (25677)

16 "false positiv$".ti,ab [Limit to: Publication Year 2008-2015] (14575)

17 "false negativ$".ti,ab [Limit to: Publication Year 2008-2015] (7448)

18 accuracy.ti,ab [Limit to: Publication Year 2008-2015] (112597)

19 11 or 12 or 13 or 14 or 15 (1103765)

20 12 and 13 and 19 (94)

****

Screening test

21 NEONATAL SCREENING/ [Limit to: Publication Year 2008-2015] (3104)

12 screening.ti,ab [Limit to: Publication Year 2008-2015] (114395)

23 18 or 19 (145165)

24 4 and 23 (182)

****

Combined

22 9 or 20 or 24 (349)

Similar searches were also carried out in Embase and the Cochrane Library.

The above search strategies retrieved 812 references in total, after duplicate references were removed. The title and abstracts of the remaining citations were sifted for relevance to screening for CAH; 122 references were deemed to be relevant.

A second title and abstract sift of these 122 references was conducted; 55 studies were selected for appraisal at full text, 34 of which were selected for inclusion in the review.

Quality

Studies not in English, conference abstracts, non-systematic reviews, editorials, other opinion pieces, and those with nonhuman data were excluded. Case series and experimental studies of fewer than three patients were excluded except where they reported cases missed by screening. Additional relevant references identified during the preparation of the report were also included.

References

Al Hosani., et al. (2014). "Expanding the comprehensive national neonatal screening programme in the United Arab Emirates from 1995 to 2011." Eastern Mediterranean Health Journal **20**(1): 17-23.

Barra, C. B., et al. (2012). "Neonatal screening for congenital adrenal hyperplasia." Revista Da Associacao Medica Brasileira **58**(4): 459-464.

Chan, C. L., et al. (2013). "Congenital adrenal hyperplasia and the second newborn screen." Journal of Pediatrics **163**(1): 109-113.e101.

Coulm, B., et al. (2012). "Efficiency of neonatal screening for congenital adrenal hyperplasia due to 21-hydroxylase deficiency in children born in mainland France between 1996 and 2003." Archives of Pediatrics & Adolescent Medicine **166**(2): 113-120.

Dhillon, K., et al. (2011). "An automated method on analysis of blood steroids using liquid chromatography tandem mass spectrometry: application to population screening for congenital adrenal hyperplasia in newborns." Clinica Chimica Acta **412**(23-24): 2076-2084.

Dumic, K., et al. (2009). "Classical congenital adrenal hyperplasia due to 21-hydroxylase deficiency in croatia between 1995 and 2006." Hormone Research **72**(5): 310-314.

Fingerhut, R. (2009). "False positive rate in newborn screening for congenital adrenal hyperplasia (CAH)-ether extraction reveals two distinct reasons for elevated 17alpha-hydroxyprogesterone (17-OHP) values." Steroids **74**(8): 662-665.

Gidlof, S., et al. (2013). "One hundred years of congenital adrenal hyperplasia in Sweden: a retrospective, population-based cohort study." The Lancet Diabetes & Endocrinology **1**(1): 35-42.

Gidlof S., et al. (2014). "Nationwide neonatal screening for congenital adrenal hyperplasia in sweden: a 26-year longitudinal prospective population-based study. " JAMA Pediatrics **168**(6): 567-74.

Gonzalez, E. C., et al. (2013). "Newborn screening for congenital adrenal hyperplasia in Cuba: six years of experience." Clinica Chimica Acta **421**: 73-78.

Hayashi, G., et al. (2011). "Weight-adjusted neonatal 17OH-progesterone cutoff levels improve the efficiency of newborn screening for congenital adrenal hyperplasia." Arquivos Brasileiros de Endocrinologia e Metabologia **55**(8): 632-637.

Hertzberg, V. S., et al. (2011). "Birth prevalence rates of newborn screening disorders in relation to screening practices in the United States." Journal of Pediatrics **159**(4): 555-560.

Hird, B. E., et al. (2014). "No evidence of an increase in early infant mortality from congenital adrenal hyperplasia in the absence of screening." Archives of Disease in Childhood **99**(2): 158-

164.

Huynh, T., et al. (2009). "The Clinical and Biochemical Spectrum of Congenital Adrenal Hyperplasia Secondary to 21-Hydroxylase Deficiency." The Clinical Biochemist Reviews **30**(2): 75-86

Janzen, N., et al. (2011). "Rapid steroid hormone quantification for congenital adrenal hyperplasia (CAH) in dried blood spots using UPLC liquid chromatography-tandem mass spectrometry." Steroids **76**(13): 1437-1442.

Janzen, N., et al. (2012). "Neonatal screening: identification of children with 11beta-hydroxylase deficiency by second-tier testing." Hormone research in pædiatrics **77**(3): 195-199.

Khalid, J. M., et al. (2010). "Newborn Screening for Congenital Adrenal Hyperplasia: An appraisal against National Screening Committee Criteria." Unpublished

Khalid, J. M., et al. (2012). "Incidence and clinical features of congenital adrenal hyperplasia in Great Britain." Archives of Disease in Childhood **97**(2): 101-106.

Knowles, R. L., et al. (2014). "Late clinical presentation of congenital adrenal hyperplasia in older children: findings from national paediatric surveillance." Archives of Disease in Childhood **99**(1): 30-34.

Lee, J. E., et al. (2008). "Corrected 17-alpha-hydroxyprogesterone values adjusted by a scoring system for screening congenital adrenal hyperplasia in premature infants." Annals of Clinical & Laboratory Science **38**(3): 235-240.

Liivak, K., et al. (2008). "Incidence of classical 21-hydroxylase deficiency and distribution of CYP21A2 mutations in Estonia." Hormone Research **69**(4): 227-232.

Loeber, J. G., et al. (2012). "Newborn screening programmes in Europe; arguments and efforts regarding harmonization. Part 1 - From blood spot to screening result." Journal of Inherited Metabolic Disease **35**(4): 603-611.

Morikawa, S., et al. (2014). "Results from 28 years of newborn screening for congenital adrenal hyperplasia in Sapporo." Clinical Pediatric Endocrinology **23**(2): 35-43.

Nascimento, M. L., et al. (2014). "Ten-year evaluation of a Neonatal Screening Program for congenital adrenal hyperplasia." Arquivos Brasileiros de Endocrinologia e Metabologia **58**(7): 765-771.

Pezzuti, I. L., et al. (2014). "A three-year follow-up of congenital adrenal hyperplasia newborn screening." Jornal de Pediatria **90**(3): 300-307.

Rossi, C., et al. (2011). "Confirmation of congenital adrenal hyperplasia by adrenal steroid profiling of filter paper dried blood samples using ultra-performance liquid chromatography-tandem mass spectrometry." Clinical Chemistry & Laboratory Medicine **49**(4): 677-684.

Salter, S. J., et al. (2015). "Analysis of 17 alpha-hydroxyprogesterone in bloodspots by liquid chromatography tandem mass spectrometry." Annals of Clinical Biochemistry **52**(1): 126-134.

Sarafoglou, K., et al. (2012a). "Cases of congenital adrenal hyperplasia missed by newborn screening in Minnesota." JAMA **307**(22): 2371-2374.

Sarafoglou, K., et al. (2012b). "Comparison of one-tier and two-tier newborn screening metrics for congenital adrenal hyperplasia." Pediatrics **130**(5): e1261-1268.

Schreiner, F., et al. (2008). "False negative 17-hydroxyprogesterone screening in children with classical congenital adrenal hyperplasia." European Journal of Pediatrics **167**(4): 479-481.

Seo, J. Y., et al. (2014). "Steroid profiling for congenital adrenal hyperplasia by tandem mass spectrometry as a second-tier test reduces follow-up burdens in a tertiary care hospital: a retrospective and prospective evaluation." Journal of Perinatal Medicine **42**(1): 121-127.

Sharma, R., et al. (2014) "Congenital adrenal hyperplasia: issues in diagnosis and treatment in children." Indian Journal of Pediatrics 81(2): 178-185

Speiser, P. W., et al. (2010). "Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline." Journal of Clinical Endocrinology & Metabolism **95**(9): 4133-4160.

Votava, F., et al. (2012). "Lessons learned from 5 years of newborn screening for congenital adrenal hyperplasia in the Czech Republic: 17-hydroxyprogesterone, genotypes, and screening performance." European Journal of Pediatrics **171**(6): 935-940.