Four Nations Diabetic Retinopathy Screening Intervals Project Study Group
Report on Findings – *FINAL REPORT 11/12/13*

1. Executive summary

A Study Group, whose membership is given in Appendix A, was formed to address the study question: **can we determine optimal screening intervals for different risk groups that can be identified in the current data set?** We collected and analysed data from seven programmes (whole nation programmes in Scotland, Wales and Northern Ireland, and a sample of four of the 84 English programmes). This report presents a summary of findings from the Study Group and a brief synopsis of the evidence on which these are based. Further detail on the analyses undertaken and research reviewed is given in a presentation linked to this document.

The Study Group's **overriding conclusions** are as follows:

- **If accurate and consistent grading were assured, the data from the seven programmes suggest that an appropriate yield for identifying diabetic retinopathy in screening would be 2.5%, at which point the optimal intervals would be two to three years for the low risk group, annual for medium risk, and six monthly for the high risk group[1].** (The definitions of these risk groups are explained later in this report).

- **Robust systems of internal and external Quality Assurance are necessary to meet the requirement of accurate and consistent grading.**

- **Our analyses demonstrate a range of issues to be addressed in any future implementation of risk-based variable intervals, including tackling unwarranted variation in grading practice, ensuring better capture of basic patient-level data and the need for prospective analyses to track the impact on outcomes.**

The Study Group's **supporting conclusions** are as follows:

- Findings from new analyses conducted for this study and the outcome of the linked literature review consistently support the view that there is a sizeable group of people with diabetes who have no retinopathy or mild background retinopathy (in one eye only at two consecutive screening episodes) who are at low risk of progression to STDR and could be screened less often.
- Accurate and consistent grading is a prerequisite for implementing risk-based intervals. Our analyses have identified unwarranted variation in grading practice which results in differences across programmes in the likelihood of patients being assessed as having DR and referral for onward treatment. The risk is that programmes with grading centres which identify fewer patients as having DR than suggested by comparative data may be insensitive. In these programmes patients may be inappropriately assigned to a low risk group. Our analyses across the Four Nations suggest that variation is associated with differences in grading rather than differences between the programme populations.

[1] The study approach segments patients into nine risk groups based on screening results at two consecutive screening episodes. Our analyses concentrate on Low risk (Group 9: R0 in both eyes on first occasion and second occasion) and High risk (Group 1: R1/M0 in both eyes on first occasion and on second occasion). There are seven intermediate groups, and the Medium referred to in this report is Group 5 (R1/M0 in one eye only and R0 in the other eye, on first occasion and second occasion).

- Our analyses show that patients who have had diabetes for a long time before their first screen are highly likely to have DR. If programmes were to identify and screen such patients earlier then much avoidable vision loss could be prevented.
- The task of identifying these patients and those in other risk groups is hampered by the unacceptable failure to comprehensively and systematically capture basic data such as date of diagnosis with diabetes and gender. Further weaknesses in the current dataset captured by screening programmes and the wider NHS severely limit understanding of the effectiveness and impact of current practice, such as measuring patient outcomes or the impact of screening for different ethnic groups.

Limitations
Limitations to our analyses include the selection of English programmes, as our study includes four of 84 programmes, which agreed to provide data and met pre-specified inclusion criteria. Our findings cannot be generalised to other English programmes and further analyses of each programme's data would be required before definitive conclusions could be drawn. Our results are based on grading as recorded on relevant software; there may be missing data or activity not recorded. Our study did not include a cost effectiveness analysis.

Wider considerations
The Study Group considers that a new RCT would bring new knowledge and fill some current gaps, such as investigating the link between screening intervals and long term progression of eye disease. Discussions around existing evidence and the potential for a new RCT have raised the question of what is the appropriate level of evidence to assess public health interventions. The standard hierarchy of evidence, with its emphasis on experimental rather than observational data (which favours prospective, randomised, trials with comparison groups) may not be an appropriate yardstick with which to measure the real-world experience of screening programmes.

Recommendation
Based on these findings the Study Group recommends to the Four Nations Steering Group that our conclusions are shared with the UK National Steering Committee and that the issues identified through our work are addressed urgently by the relevant Four Nations Programmes.

The remainder of this report sets out: the context and process for our research; findings from new analyses undertaken; brief summary of an associated rapid review of the literature; conclusions and recommendations; next steps suggested by the Study Group. A presentation 'Supporting Information' is available alongside this report and gives detailed outcomes of the analyses conducted. Also, each participating programme has received a pack of information containing its specific results.

2. Context for the research

About 5% of the UK population has a diagnosis of diabetes. Most (~90%) have Type 2 diabetes (T2DM) and numbers with T2DM are rising because of population changes including increasing longevity and obesity, low levels of physical activity and an altering ethnic mix. The proportion of those with T2DM identified is also growing because of opportunistic screening, introduction of 'health checks' and changes in diagnostic criteria. Levels of Type 1 DM are also rising by ~ 5% per annum, for reasons which remain unknown. Annual screening for DR in people with diabetes is recommended for all those with diabetes aged 12 and above. Those found to have STDR are referred on for further assessment and for treatment if required. Treatment with laser therapy has been shown to reduce the risk of vision loss. The introduction of new and clinically effective

therapies for diabetic maculopathy means that it is important to identify those at risk of this complication.

Due to the increasing numbers of people needing to be screened, it has become important to investigate whether it is possible to stratify patients into risk categories to determine appropriate screening intervals. Those at low risk might be able to have screening intervals set at two or more years without an overall elevated risk of missed detection of STDR, and those at high risk might be invited at six-monthly intervals. Also, greater effort and resources could be put into encouraging and enabling patients who have never been screened to attend.

It is known that it is possible to stratify patients with diabetes into levels of risk using clinical information (HbA1c, duration and type of diabetes, blood pressure and previous screening results). However, these items of information are not always known to screening programmes. Date of diagnosis of diabetes has not been a mandatory item of information and there is currently no routine collection of clinical information from primary care by screening programmes.

Gloucestershire Diabetic Retinopathy Research Group developed a model using level of DR at each of two consecutive screening episodes to stratify groups of patients at high and at low risk of progression to STDR (R2 or R3 or M1 or any combination of these) from no diabetic retinopathy (No DR or R0) or from background DR (also called Mild Non Proliferative DR, NPDR). STDR is the level at which patients are referred from the screening programme to referral centres or to a hospital eye service). This approach allows patients to be allocated to one of nine risk groups, given in Appendix B. The approach developed in the Gloucester study has provided the method for this research.

3. Process – how we went about our work

We recruited data for the study from seven DR screening programmes: whole nation programmes in Wales, Scotland and Northern Ireland, and four English programmes: Brighton, Derbyshire, Leeds and Staffordshire. Inclusion criteria for the English programmes included: voluntary participation; minimum 10,000 patients in 2005; and having no Serious Untoward Incident relating to grading. We also aimed for a representative sample of programmes, encouraging those not usually involved in such research and with variation in socio-demographic measures and population ethnicity. However, with only four of 84 English programmes included, caution must be taken in generalising results.

At the outset the approach was referred to the Chair of a Research Ethics Committee who confirmed that formal approval was not required, as we were using existing, routinely collected clinical data and all analyses would be fully anonymised.

A data set was defined comprising core demographic information for each patient (anonymised) and a linked file of screening episode results. Data were sent by participating programmes to Yorkshire and Humber Public Health Observatory (now part of Public Health England), who reviewed, cleaned and prepared the data for analysis by Irene Stratton, lead statistician. Arrangements were made for security and confidentiality in data transfer. The Study Group was consulted on the definition of this dataset and in the subsequent analysis and interpretation, by regular e-mails and at its three meetings (26 November 2012; 12 March 2013; 5 September 2013).

Criteria for the data to be used in the analysis included having:

- Screening and grading results after 1/1/2005
- At least 3 grading episodes with fully graded images of both eyes
- First two episodes with no referable DR

It was known that the grading protocols differed across the Four Nations (hence there was never the option to simply pool all the UK data). Also, grading protocols were not identical across English programmes (e.g. grading of 1 microaneurysm, arbitration of R0/R1M0 discrepancies). An exercise to map grading protocols was undertaken to address these points. Our analyses were based on the data recorded on relevant software; there may be missing data or activity not recorded.

The study has, for the first time, combined a large pool of data from seven UK screening programmes. Around 354,000 patients met inclusion criteria resulting in approximately 1.6 million screening episodes being used in the analyses. The median number of screening episodes per person included was five.

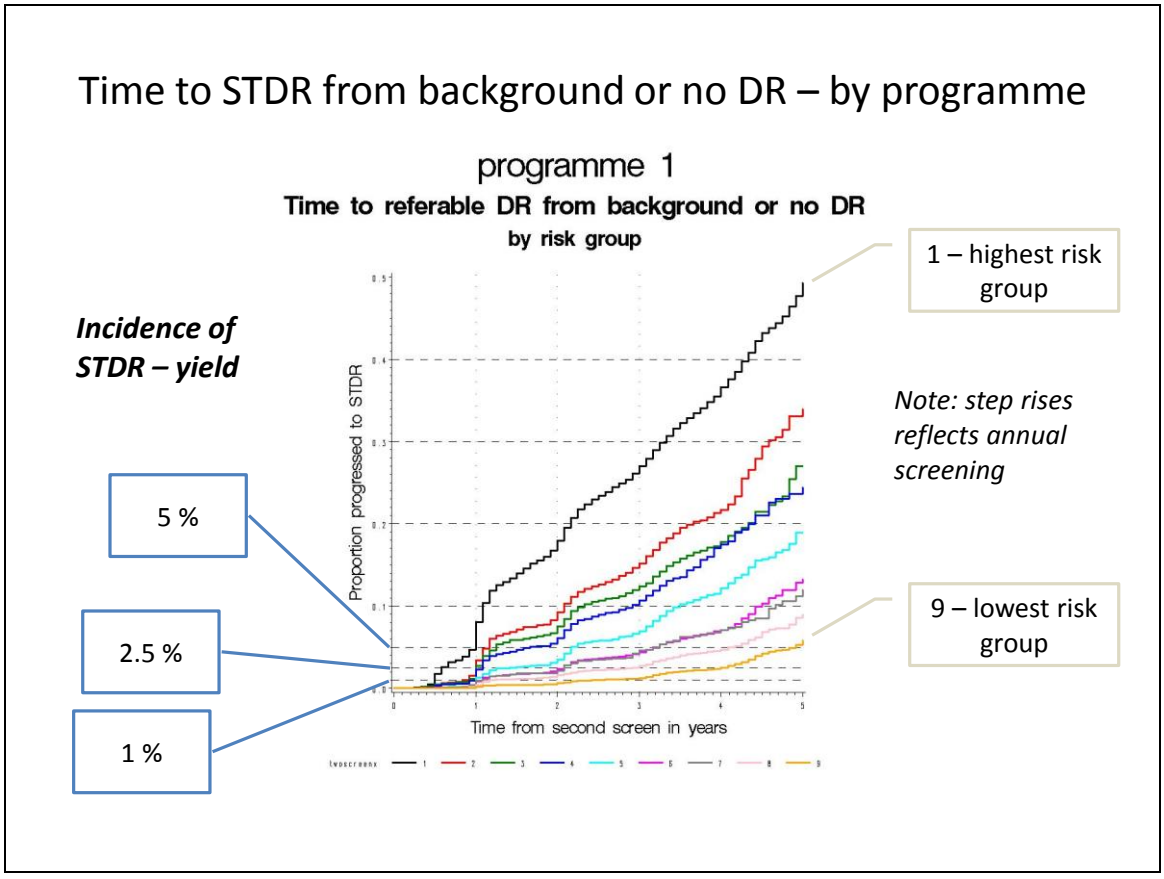
4. Findings from our analyses

Our results are set out under the following seven headings given as questions to be addressed.

4.1. *What is the overall time for progression to STDR for each of the seven programmes?*

Analysis was undertaken for each of the seven programmes to determine the progression to STDR across the nine risk groups. At this point it is important to introduce the concept of 'yield' in screening, in this case the incidence of STDR identified for referral. We note that there is no consensus in the literature or clinical practice on what the ideal or standard level of yield should be (or what percentage of patients would be expected to progress for referral). We make an assumption that the yield at a level of 2.5% is lower than the current practice and is a reasonable, conservative approach.

The graph below presents the outcome of this analysis for Programme 1, as an illustration.



Findings and interpretation

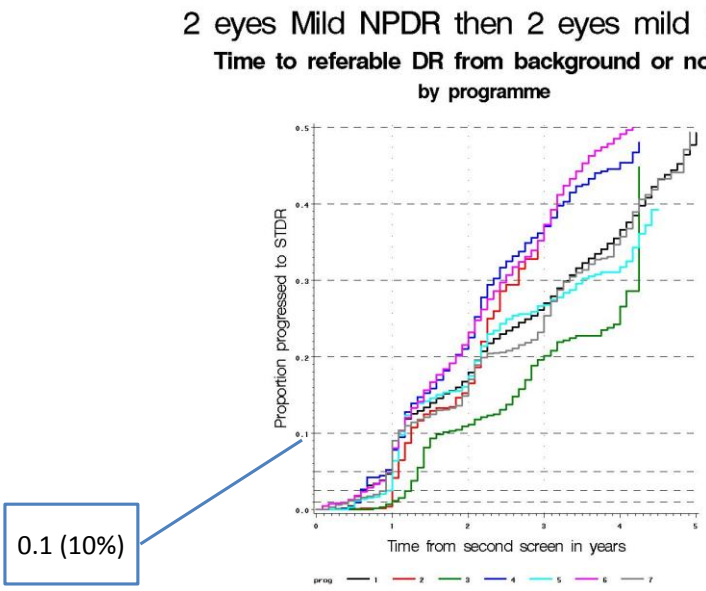
- The algorithm appropriately identifies different levels of risk. For example, it shows that the highest risk group, those with R1 in both eyes at two screens (R1/R1), most rapidly progresses to STDR, while the lowest risk group R0/R0 has the lowest progression.
- The ranking of risk groups varies across programmes and differs from that in Gloucester, in that it is less distinct.
- Our methods are appropriate for addressing what would be an appropriate interval.

4.2. *What is the time for progression to STDR for three selected risk groups in each of the seven programmes?*

Analysis was undertaken for three risk groups: high risk (R1/R1), low risk (R0/R0) and medium risk (all other seven risk groups). Results for all seven participating programmes have been plotted together. The question was raised: how many patients fall into the 'low risk' group? The data indicate that the proportion of image sets graded as R0/R0 from two graded sets in those without STDR varies across programmes in these analyses from 50% to 79%, with a median of 64%. So, around one-half of patients without referable retinopathy are in the lowest risk group.

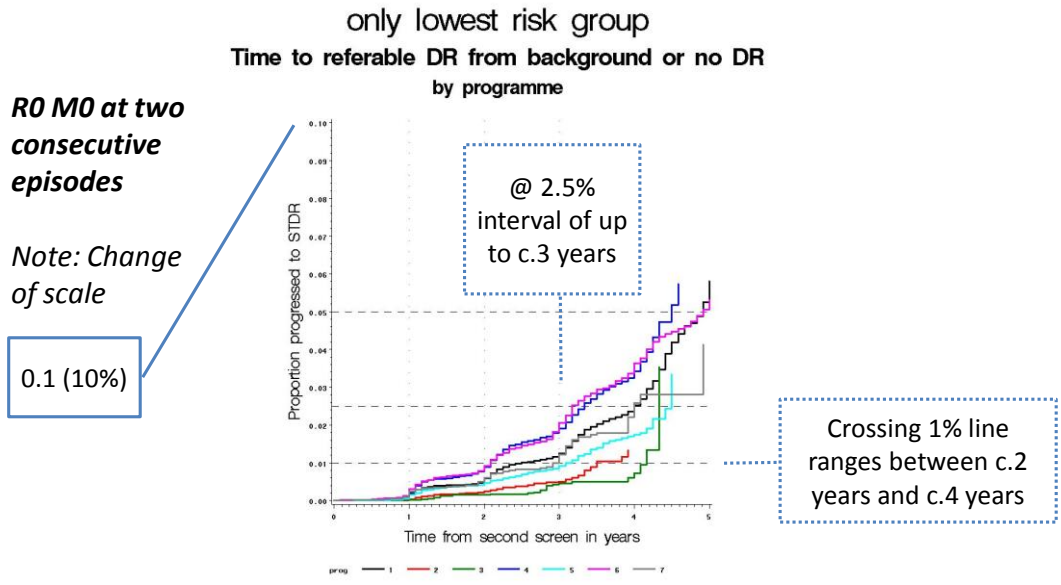
The graphs for the high and low risk analyses are given below:

Time to STDR for all seven programmes – by three risks groups: **HIGH**



One way of explaining the above is that in one year after the second screen, around one in ten of the high risk group will need to be referred (STDR), and after five years, half this group will be referred.

Time to STDR for all seven programmes – by three risks groups: **LOW**



Findings and interpretation

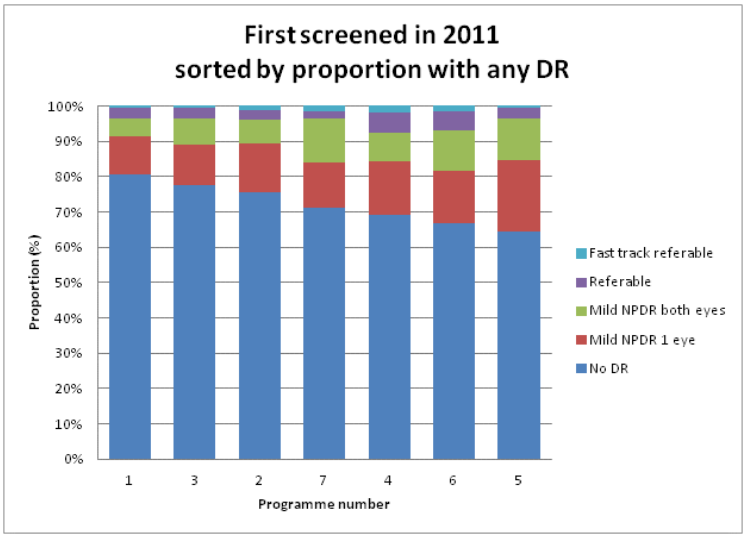
- For the lowest risk group (RO/MO, twice), for patients included in the study, the interval could be extended to two years without increasing the risk of undetected STDR and associated sight loss.
- Middle risk stay at 1 year.
- Highest risk could reduce to 6 months as the overall risk is ~10% at 1 year.

However, this analysis and that summarised in 4.1 identifies marked variation which implies either: populations differ significantly between programmes or grading differs significantly between programmes. For example, the time to reach a given yield in both high and low risk groups varies across the seven programmes in the previous two graphs. Detailed consideration of a wide range of analyses undertaken for this study and other supporting data indicate that variation in grading is the key factor in the different results, as illustrated below.

4.3. *How does the number of people graded as having DR vary by programme, for people at first screen in 2011?*

The graph below shows the proportion of patients who were screened for the first time in 2011 found to have DR across the seven programmes:

Grading at first screen – 2011, by programme



All patients first screened in 2011, for all programmes – proportion graded with DR ranges from c.20% to c.35%

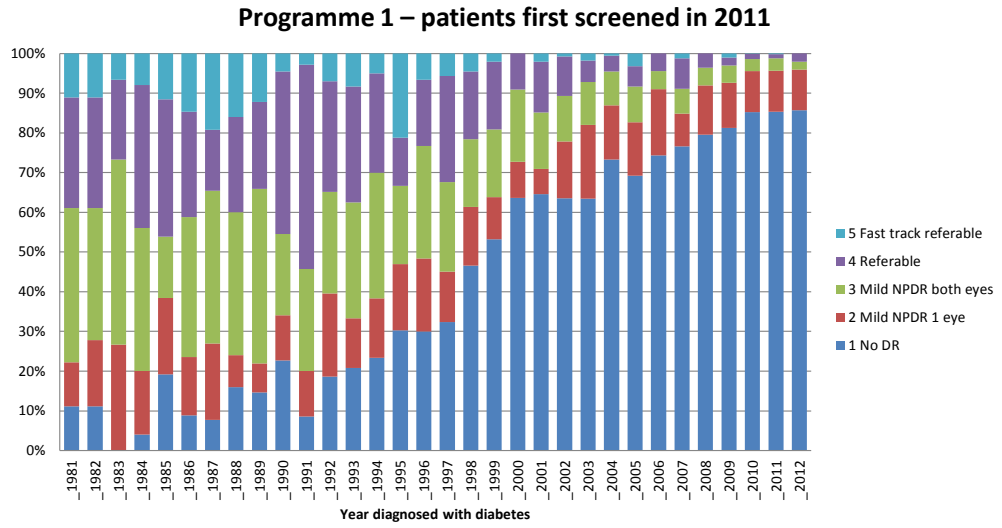
Findings and interpretation

- The graph above indicates that the proportion of patients first screened in 2011 found to have any DR varies from ~20% (programme 1) to ~35% (programme 5).
- Investigation of the possible reasons for this variation, such as differences in the detection or incidence of diabetes across the programme populations, led to the following analysis.

4.4. *How does the number of people graded as having DR vary by time since year of diagnosis with diabetes?*

The graph below shows, for Programme 1, the proportion of patients first screened in 2011 who were found to have DR separated by their year of diagnosis with diabetes. The results for other programmes are similar, where data were available (i.e. date of diagnosis with DM).

Grading by year of diagnosis – one example



Findings and interpretation

- While most patients are screened relatively soon after diagnosis with diabetes, this striking graph shows that, in Programme 1, people diagnosed with diabetes before 1990 and first screened in 2011 had ~90% likelihood of being assessed as having some form of DR.
- For the programmes in this study and people with known dates of diagnosis, 39,254 patients were first screened in 2011; of these 1,730 had referable DR (4.4%), and 384 (1%) had R3.
    - Of 813 people diagnosed with diabetes between 1980 and 1989 and first screened in 2011, (27%) people had referable retinopathy and of these 71 (9% of those screened) were graded with fast-track referable retinopathy (R3).
    - Of 2,861 diagnosed between 1990 and 1999 and first screened in 2011, 471 (16.5%) had referable retinopathy and 116 (4%) had R3.
    - So half the R3 in those first screened in 2011 was in patients diagnosed before 2000.
- The analysis demonstrates the importance of identifying people who have been diagnosed with diabetes but have never been screened, to get them screened at least once and help prevent avoidable vision loss.

To compare across programmes, we removed the confounding factor of duration of known diabetes and limited the analysis to people screened within 1 year of diagnosis, explained below.

4.5. *How does the number of people graded as having DR vary by Programme over time, for people screened within a year of diagnosis?*

Analysis was conducted which illustrates that programmes vary over time in the proportion of those screened within one year of diagnosis assessed as having DR – some are increasing, some stay the same, and other programmes are identifying a smaller proportion. (The graphs are given in the 'Supporting Information' presentation available alongside this report.)

Findings and interpretation

Variation in results for this and previous analyses both between programmes and within programmes over time cannot be fully explained based on the current dataset. Three main areas where there may be systematic difference that contributes to the variation are as follows:

- Demographics (e.g. age, socioeconomic, ethnicity)
- Programme level (e.g. attendance, protocols, practice)
- Grading (e.g. training, workforce, techniques, testing)

Further work has been undertaken to investigate variation in grading by geographic area and grading centre.

4.6. *What is the further evidence of variation in grading across the Four Nations?*

Detailed analyses have been undertaken for English programmes using a separate dataset from DESP annual returns. Further analyses have been undertaken for Scotland, Wales and Northern Ireland using our dataset. Graphs are presented in the associated presentation.

Findings and interpretation

- For England, the 84 programmes have unwarranted[2], unexplained variation in the proportion of people graded and found to have DR. This is the case for various sub-groups, i.e. for the detection of retinopathy, maculopathy, STDR and 'fast track' DR. We interpret this to be primarily an issue of quality and consistency in grading practice and note that some centres appear to be too sensitive (identifying more than expected) and others insensitive (identifying fewer than expected).
- For Scotland, there is some evidence of variation across the nine grading centres (note: Scotland has 14 Health Boards and five regional screening centres, some of which have more than one grading centre).
- The least amount of variation is found in Wales and Northern Ireland, both of which have single grading centres. Analysis by health board (seven in Wales, four in Northern Ireland) suggests that there is no significant difference in rates of progression despite varying demographics. This provides further support for our contention that the primary issue is grading practice.
- It is for each programme to address the results of these findings. We argue that robust systems of internal and external Quality Assurance are necessary to ensure grading is accurate and consistent. For example, an array of testing methods could be deployed to reduce variation, including use of: masked image sets; grading by grader; comparison across programmes; Test and Training data (monthly test sets should be undertaken by all graders with results reported to grader and screening lead); and use of external reference graders.

4.7. Additional analyses

A range of other sub-questions were investigated in our study, of which we highlight the following two:

a. Separation of the group of patients identified as having STDR, i.e. referred to hospital eye service, between those identified with R2/R3 and M1/M2. The results indicate there were

[2] The term 'unwarranted' draws on the work of Professor John E Wennberg (whose approach in the Dartmouth Atlas of Health Care has been adapted in the NHS) and refers to variation which is not explained by illness, need or patient preference. See: BMJ 2011; 342:d1513.

significant differences between programmes by risk group and the problem was apparent in both R2/R3 and maculopathy grading.

b. Analysis of how quickly patients are screened after the date on which they are added to the screening register, and what proportion are ever screened. The results show that, for the programmes which provided the necessary data (n=4), between 5% and 10% of patients are not screened within three years (of being added to the screening register). Further analysis suggests that the oldest patients (over 85 years) and those in the range 18 to 24 and 25 to 34 years are particularly at risk of not being screened.

5. Findings from the literature review

A team at the University of Warwick Medical School, led by Professor Aileen Clarke, were commissioned by the Four Nations Steering Group to undertake two rapid reviews of the literature. The primary review addressed the question: **Would changing diabetic eye screening intervals from the current annual recommendation lead to changed clinical outcomes?** A summary of the outcome from the literature review is given below. A supplementary search sought to address the question: **Does changing screening intervals in any screening programme alter uptake?** No literature or evidence was found which met the inclusion criteria for this second question, so no conclusions could be reached about the potential impact on uptake (or attendance) of changing screening intervals.

Brief Summary of the Warwick Rapid Literature Review

Aim: To determine if changes in screening intervals for diabetic eye screening from the currently recommended annual screening would lead to changed clinical outcomes
Objectives: To perform a systematic review of the current literature; to critically appraise the identified current literature; to synthesise the findings of the literature search in a narrative format.
Methods: The major medical literature databases were searched using standard methods.

Results: From 12,063 titles / abstracts, 129 publications were evaluated at full text level, from which 25 studies met the inclusion criteria, comprising:

- 10 observational studies of existing screening programmes or participants in ongoing trials;
- 10 economic analyses (n= 10) all of which were based on modelling;
- 5 studies describing the development/evaluation of risk stratification algorithms

It is noted that:

- The majority of participants in the identified studies had T2DM and had no background diabetic retinopathy at baseline.
- Most of the observational studies of existing programmes reported clearly formulated objectives, population characteristics, main outcome measures and had adequate methods of participant recruitment, exposure measurement (i.e., types of screening tests and between-test intervals), and outcome ascertainment.

All the identified observational studies identified concluded that in low risk patients the screening interval could safely be extended to beyond one year, with a number of caveats. The definition of low risk patients varied, depending on factors measured and included controlled diabetes on dietary treatment, controlled blood pressure and duration of diabetes of less than 10 years. The evidence from cost-effectiveness studies was less clear, but generally supported the findings of observational studies for adopting longer screening intervals for low risk patients and suggested that biennial screening intervals could be adopted for those with no background retinopathy. Risk stratification

algorithms showed potential for safely increasing the screening interval.

Discussion:

Strengths of the evidence base include:

- Some key single cohort studies are large in size and include recent cohorts.
- Competent statistical analysis of the data and most larger studies are based on UK populations.
- Majority of cohort studies appear well conducted according to appraisal frameworks.
- Economic studies are often highly complex, covering many variables.

Weaknesses of the evidence base include:

- Lack of randomised trials comparing the effectiveness of screening programs utilising different intervals between the screens.
- None of the included studies were of a comparative nature which would allow direct comparison of progression to (or incidence) of sight-threatening retinopathy/vision loss between subjects with diabetes screened with different intervals.
- Studies use different methodologies to screen and grade retinopathy – patients reported as 'no background retinopathy' should be considered to be 'patients in whom diabetic retinopathy has not been found' due to the different sensitivities and specificities of screening and grading.
- High attrition rate in almost all studies and implications not always clear – the non attendees may differ systematically to the attendees and this is not always explored or accounted for in the studies in any great depth.
- Many studies make recommendations based on caveats relating to risk factors such as duration of diabetes, recorded HbA1c and age at onset. This data is not currently routinely collected or linked in parts of the UK as part of the screening programme.
- Further weaknesses include: older and smaller data sets used for some economic analyses and findings not reported in relation to different ethnicities or socio economic groups may mean that findings are less applicable to today's populations and lack of evidence from 'real world' data on potential cost effectiveness from the studies that used simulation models.

The authors recognise limitations of the Review, listed in the full document, and difficulties including the heterogeneous data which made a narrative approach inevitable but drawing evidence based conclusions very difficult. After due consideration, the conclusions of the Review were as follows

Although most economic modelling studies and single cohort studies point to little difference between clinical outcomes from annual and biennial screening programmes in people in whom no evidence for background retinopathy has been found, no real world observational or randomised comparisons exist and these conclusions are associated with considerable uncertainty that is difficult to gauge adequately.

A direct comparison study is required. Such research would also underpin a rigorous cost effectiveness analysis and would provide reliable evidence about the suggested link between longer screening intervals and compliance with scheduled visits. The available evidence is inadequate to fully inform a policy decision.

In the absence of a properly conducted trial, the use of individual patient data from an appropriate UK-based screening programme in conjunction with individual patient simulation and bottom up costing of the programme would provide an improved evidence base than is currently available.

It is noted that two additional studies were identified after the Review was completed. The first, a US Review '*Screening intervals for diabetic retinopathy and incidence of visual loss: a systematic review*' (Echouffo-Tcheugui et al, Diabetic Medicine, July 2013), also included 25 studies and

concluded that: *A 2-year screening interval for people with no sight threatening diabetic retinopathy at diagnosis may be safely adopted. For patients with pre-existing diabetic retinopathy, a shorter interval of ≤ 1 year is warranted.* We note that this review does not consider the issue of grading. The second paper, 'Predicted impact of extending the screening interval for diabetic retinopathy: the Scottish Diabetic Retinopathy Screening programme' (Looker et al, Diabetologia, Aug 2013), concluded: *Transition rates to referable diabetic eye disease were lowest among people with type 2 diabetes and two consecutive screens showing no visible retinopathy. If such people had been offered two yearly screening the DRS service would have needed to screen 40% fewer people in 2009.*

6. Conclusions and recommendations

The Study Group's **overriding conclusions** are as follows.

- **If accurate and consistent grading were assured, the data from the seven programmes suggest that an appropriate yield for identifying diabetic retinopathy in screening would be 2.5%, at which point the optimal intervals would be two to three years for the low risk group, annual for medium risk, and six monthly for the high risk group.**

- **Robust systems of internal and external Quality Assurance are necessary to meet the requirement of accurate and consistent grading.**

- **Our analyses demonstrate a range of issues to be addressed in any future implementation of risk-based variable intervals, including tackling unwarranted variation in grading practice, ensuring better capture of basic patient-level data and the need for prospective analyses to track the impact on outcomes.**

The Study Group's **supporting conclusions** are as follows.

- Findings from new analyses conducted for this study and the outcome of the linked literature review consistently support the view that there is a sizeable group of people with diabetes who have no retinopathy or mild background retinopathy (in one eye only at two consecutive screening episodes) who are at low risk of progression to STDR and could be screened less often.
- Accurate and consistent grading is a prerequisite for implementing risk-based intervals. Our analyses have identified unwarranted variation in grading practice which results in differences across programmes in the likelihood of patients being assessed as having DR and referral for onward treatment. The risk is that programmes with grading centres which identify fewer patients as having DR than suggested by comparative data may be insensitive. In these programmes patients may be inappropriately assigned to a low risk group. Our analyses across the Four Nations suggest that variation is associated with differences in grading rather than differences between the programme populations.
- Our analyses show that patients who have had diabetes for a long time before their first screen are highly likely to have DR. If programmes were to identify and screen such patients earlier then much avoidable vision loss could be prevented.
- The task of identifying these patients and those in other risk groups is hampered by the unacceptable failure to comprehensively and systematically capture basic data such as date of diagnosis with diabetes and gender. Further weaknesses in the current dataset captured by screening programmes and the wider NHS severely limit understanding of the effectiveness and impact of current practice, such as measuring patient outcomes or the impact of screening for different ethnic groups.

Limitations
Limitations to our analyses include the selection of English programmes, as our study includes four of 84 programmes, which agreed to provide data and met pre-specified inclusion criteria. Our findings cannot be generalised to other English programmes and further analyses of each programme's data would be required before definitive conclusions could be drawn. Our results are based on grading as recorded on relevant software; there may be missing data or activity not recorded. Our study did not include a cost effectiveness analysis.

Wider considerations
The Study Group considers that a new RCT would bring new knowledge and fill some current gaps, such as investigating the link between screening intervals and long term progression of eye disease. Discussions around existing evidence and the potential for a new RCT have raised the question of what is the appropriate level of evidence to assess public health interventions. The standard hierarchy of evidence, with its emphasis on experimental rather than observational data (which favours prospective, randomised, trials with comparison groups) may not be an appropriate yardstick with which to measure the real-world experience of screening programmes.

Recommendation
Based on these findings the Study Group recommends to the Four Nations Steering Group that our conclusions are shared with the UK National Steering Committee and that the issues identified through our work are addressed urgently by the relevant Four Nations Programmes.

7. Proposed next steps

Members of the Study Group, in co-operation with the Steering Group, intend to write a series of short papers based on the new analyses undertaken for this study, which will be submitted for publication in peer reviewed journals. While this report concludes the work of the Study Group, members would be pleased to discuss the potential to further support the Four Nations Steering Group and UK National Steering Committee.

Appendix A – Study Group Membership

Professor Brian Ferguson, Director for Knowledge & Intelligence (England), Public Health England (Chair of the Study Group)

Professor Max Bachmann, Professor of Health Services Research, Norwich Medical School, University of East Anglia

Dr Daniel Chalk, Associate Research Fellow, University of Exeter

Mr Colin Jones, Consultant Ophthalmologist, Norfolk and Norwich University Hospitals NHS Foundation Trust

Martin Land, Director, Landmark Health Consulting (Project Manager for the Study)

Professor Graham Leese, Consultant and Professor in Diabetes and Endocrinology, Ninewells Hospital and School of Medicine, University of Dundee

Dr Helen Looker, Clinical Senior Research Fellow, School of Medicine, University of Dundee

Professor Sue Moss, Professor of Cancer Epidemiology, Centre for Cancer Prevention, Queen Mary University of London

Irene Stratton, Honorary Associate Professor, University of Warwick Clinical Sciences Research Institute and Senior Statistician, Gloucester Diabetic Retinopathy Research Group (Lead Statistician for the Study)

Appendix B – Grading Definitions and Risk Groups

R0	identifies no detected diabetic retinopathy (equivalent to ETDRS level 10).
R1(mild NPDR or background DR)	a minimum of at least the presence of one microaneurysm and/or retinal haemorrhage, equivalent to ETDRS levels 14-35
R2 (moderate to severe NPDR or pre-proliferative DR)	presence of multiple deep, round or blot haemorrhages and/or definite intraretinal microvascular abnormality (IRMA) and/or venous beading and/or reduplication, equivalent to levels 43 - 53 on the ETDRS scale
R3(proliferative DR).	presence of proliferative diabetic retinopathy (including fibrous proliferation), equivalent to a minimum of ETDRS level 61
M0	Complement of M1
M1(maculopathy)	presence of 2-dimensional photographic markers of diabetic maculopathy, specifically exudate within 1 disc diameter (DD) of the centre of the fovea, circinate or group of exudates within the macula or any microaneurysm or haemorrhage within 1DD of the centre of the fovea but only if associated with a best VA of worse than 0.3 logMAR (equivalent to Snellen 6/12).

Risk level	First screen	Second screen
1	R1 both eyes	R1 both eyes
2	R1 one eye	R1 both eyes
3	R0 both eyes	R1 both eyes
4	R1 both eyes	R1 one eye
5	R1 one eye	R1 one eye
6	R0 both eyes	R1 one eye
7	R1 both eyes	R0 both eyes
8	R1 one eye	R0 both eyes
9	R0 both eyes	R0 both eyes