Diabetologia Springer

Safety, efficacy and cost effectiveness of individualised screening for diabetic retinopathy: the ISDR open-label, equivalence randomised controlled trial

Journal:	*Diabetologia*
Manuscript ID	Draft
Manuscript Type:	Article
Keywords:	3.06.04 Retinopathy, 3.06 Complications (all), 1.02 Clinical science

Safety, efficacy and cost effectiveness of individualised screening for diabetic retinopathy: the ISDR open-label, equivalence randomised controlled trial

Deborah M Broadbent,[1] Amu Wang,[1] Christopher P Cheyne, [2, 3] Marilyn James PhD,[4] James Lathe,[4] Irene M Stratton,[5] John Roberts,[6] Tracy Moitt,[3] Jiten P Vora,[7] Mark Gabbay,[8] Marta García-Fiñana,[2,3] Simon P Harding [1] and the ISDR Study Group.

Affiliations

1. Department of Eye and Vision Science, Institute of Life Course and Medical Sciences, University of Liverpool and St Paul's Eye Unit, Royal Liverpool University Hospital, members of Liverpool Health Partners, Liverpool, UK
2. Department of Biostatistics, University of Liverpool, member of Liverpool Health Partners, UK
3. Clinical Trials Research Centre, Liverpool, UK
4. Division of Rehabilitation, Ageing and Wellbeing. School of Medicine, University of Nottingham, UK
5. Gloucestershire Retinal Research Group, Cheltenham General Hospital, Cheltenham, UK
6. Mersey Diabetes Support Group. Liverpool, UK
7. Department of Diabetes and Endocrinology, Royal Liverpool University Hospital, UK
8. Department of Health Services Research, University of Liverpool, and Brownlow Health Centre, UK, members of Liverpool Health Partners, Liverpool, UK

Correspondence:

Dr Deborah M Broadbent
Honorary Senior Lecturer,
Department of Eye and Vision Science,
3rd Floor University Clinical Departments
Royal Liverpool University Hospital
Prescot Street
Liverpool L7 8XP
Tel +44 151 706 2408
Fax +44 151 706 2409
deborah.broadbent@rlbuht.nhs.uk

Word count main manuscript 3968

Abstract

Aims:
Varying diabetic retinopathy (DR) screening intervals, informed by personal risk-levels, offers improved engagement of people with diabetes (PWD), and offers reallocation of resources to high risk groups, while addressing the increasing prevalence of diabetes. However safety data on extending intervals are minimal. We evaluated the safety, efficacy and cost effectiveness of individualised variable-interval risk-based population screening compared to usual care, with design input from PWD.

Methods:
Two-arm, parallel assignment, equivalence randomised controlled trial (minimum 2 year follow-up) in PWD aged ≥12 years registered with one English screening programme. Randomisation was 1:1 to individualised screening (6, 12 or 24 months for high, medium and low risk) determined at each episode by a risk calculation engine, using local demographic, screening and clinical data, or to annual screening (control). Primary outcome was attendance (safety). A secondary safety outcome was the development of sight threatening

DR (STDR). Cost effectiveness was evaluated within a 2 year time horizon from NHS and societal perspectives. Trial registration: ISRCTN 87561257

Results:
4534 participants were randomised. Numbers after withdrawals were: individualised arm 2097, control arm 2224. Attendance rates at first follow-up were equivalent (individualised 83.6%; control 84.7%) (difference -1.0, 95% CI -3.2 to 1.2). STDR detection rates were non-inferior: individualised 1.4%, control 1.7% (-0.3, -1.1 to 0.5). Sensitivity analyses confirmed findings. Mean differences in complete case QALYs (EQ-5D, HUI3) did not significantly differ from zero, multiple imputation supported dominance. Incremental cost savings per person were: £17.34 (CI 17.02 to 17.67), NHS perspective; £23.11 (22.73 to 23.53) societal perspective. 43.2% fewer screening appointments were required in the individualised arm.

Interpretation:
Stakeholders involved in diabetes care can be reassured by this largest ophthalmic RCT in DR screening to date that extended and individualised risk-based interval screening is feasible and can be safely and cost effectively introduced in established programmes.

Keywords
diabetic retinopathy, individualised, personalised, risk-based, screening, variable-interval

List of abbreviations

AUC	area under the curve
Chol	total cholesterol
DESP	diabetic eye screening programme
DR	diabetic retinopathy
DW	data warehouse
ESM	electronic supplementary material
GLM	generalised linear models
HUI3	Health Utilities Index Mark 3
ISDR	Individualised Screening for Diabetic Retinopathy
ITT	intention-to-treat
LCTC	Liverpool Clinical Trials Centre
LDESP	Liverpool Diabetes Eye Screening Programme
logMAR	log of the minimum angle of resolution
MI	multiple imputation
NHS	National Health Service
PP	per protocol
PWD	person with diabetes
QALY	quality adjusted life year
RCE	risk calculation engine
STDR	sight threatening diabetic retinopathy
UK	United Kingdom
VA	visual acuity

Research in context

What is already known about this topic

- Changing from annual fixed interval screening to a variable interval based on an individual's own risk will help to address increasing prevalence and improve targeting.
- Evidence is limited to a few observational studies on extension to 2-yearly screening for people with no DR, with no data on safety or acceptability of increased intervals and no randomised controlled trials.
- Cost effectiveness evidence is heterogeneous and based on modelling.

What is the key question?

- What is the safety, efficacy, feasibility and cost effectiveness of individualised variable-interval risk-based screening for sight threatening diabetic retinopathy?

What are the new findings?

- Our RCT provides strong evidence that varying intervals using an individualised risk-based approach has equivalent safety to annual fixed interval screening, with better efficacy and substantial improvements in cost-effectiveness.

How might this impact on clinical practice in the foreseeable future?

- Extended and individualised variable-interval risk-based screening can move to an implementation phase in established programmes allowing targeted reallocation of resources.

Introduction

Early detection of sight threatening diabetic retinopathy (DR, STDR) at a stage allowing timely intervention, through systematic programmes of screening, is universally recognised to be important in preventing visual impairment [1] and reducing its associated costs, but approaches vary greatly worldwide. The frequency of screening has to date been annual, based on consensus, and remains the recommendation in major guidelines.[2,3] Prevalence of diabetes is increasing rapidly,[4] and resources are stretched.

Extending the interval between screen episodes offers potential cost savings. Some developed countries have recommended or implemented 2-yearly and sometimes longer intervals for people at low risk of progression. Much evidence supporting extended intervals comes from observational studies from areas with low incidence rates,[5-8] and from modelling studies.[9] In England and Wales extended intervals have not been adopted, largely due to safety concerns highlighted by a recent systematic review calling for RCT and cost effectiveness evidence,[10] and recent failures in cancer screening.[11] In addition the feasibility of connecting large and disparate datasets is considered to be challenging.[12]

Based on our previous incidence data,[5] we designed a RCT (Individualised Screening for DR, ISDR) to investigate the safety, efficacy and cost effectiveness of extending screening intervals in low-risk PWD, with a more frequent interval for high-risk people. We utilised the emerging methodology and technologies of personalised risk prediction [13,14] and data linkage, to develop an individualised risk-based variable-interval screening approach. Individualised clinical care offers opportunities for improved patient engagement. We also wanted to test the feasibility and stability of linking routine data across varying routine NHS domains in an integrated approach. We tested the hypothesis of equivalence between the attendance rates, as a primary measure of safety, for individualised and annual screening.

Methods

Study design and participants

PWD attending for DR screening were invited to participate in a single site, two-arm parallel assignment, equivalence, RCT conducted in all community screening clinics in the Liverpool Diabetic Eye Screening Programme (DESP, LDESP), part of the English National DESP. The rationale, design and methodology have been published elsewhere,[15] and the protocol, statistical and health economics analysis plans are available online.[16] A patient and public involvement (PPI) group was embedded in all aspects of design, delivery and interpretation. The Liverpool Clinical Trials Research Centre (CTRC) developed electronic case report forms (eCRF), information systems and quality assurance. Ethical approval was by Preston NHS Research Ethics Committee (14/NW/0034).

The trial opened on May 1, 2014. Follow up was for a minimum 24 months plus a 90-day window to attend the screening invitation. Participants were recruited by trained researchers at their screening appointment, and all provided written informed consent. For children aged

12-15 years proxy consent was by the parent/guardian with, where appropriate, assent from the child. Trial management is described in the ESM.

Allocation was 1:1 to annual screening (control arm, current care) or individualised risk-based variable-interval screening recall at 6, 12 or 24 months for high-, medium- and low-risk respectively. A purpose built, dynamic data warehouse (DW) linking primary and secondary care demographic, retinopathy and systemic risk factor data populated the baseline and follow-up eCRFs (OpenClinica, LLC). Block randomisation generated by an independent statistician was conducted using a bespoke, validated electronic system in the CTRC, with stratification by clinic and age using random blocks of 4 and 6 for participants aged ≥16, and blocks of 2 for those aged <16 years to account for small numbers. Screening staff and clinical assessors were observer-masked to intervention arm, risk calculation and interval.

Procedures

Participants' risk of becoming screen positive (any of: moderate preproliferative DR, maculopathy, ungradable images, other significant sight threatening disease; full definition in ESM) was assessed by a risk calculation engine (RCE), specifically developed for the RCT.[17] Briefly, the RCE was developed using 5 years data to February 4, 2014 in the ISDR DW from 11,806 PWD. Participants and their GPs had agreed to data sharing. The RCE is a Markov multi-state model, with states defined by retinopathy level (both eyes), and transitions dependent on risk factors including historical retinopathy data. Candidate risk factors were identified in collaboration with the PPI group and selected as informative using corrected Akaike information criterion (age, diabetes duration, HbA1c, systolic BP, total cholesterol). The RCE showed good discriminatory ability (areas under the curve (AUC) 0.88, 0.90 and 0.91 for 6, 12 and 24 months, respectively).[17]

At each screening visit during the trial the RCE calculated a participant's risk of becoming screen positive using automatic exchanges of retinopathy data from the screening software (EMIS Health: OptoMize)), and risk factor data held in the DW and randomisation databases. Clinical data from primary care was updated in the DW bi-monthly. Participants were allocated to high-, medium- or low-risk against a threshold set at 2.5%, as agreed with the PPI group. The interval could change at each follow-up visit. Participants in the control arm continued with invitations to annual screening, with risk recorded for future analysis.

Participants who were screen positive attended for slit lamp biomicroscopy for determination of presence of STDR (moderate preproliferative DR or maculopathy; true positive). False positive cases were re-consented and re-entered the trial. Participants were free to withdraw consent at any time without providing a reason.

Outcomes

The primary outcome of attendance at first follow-up visit assessed the safety of individualised screening. Non-attendance was defined as failure to attend any appointment within 90 days of the follow-up invitation, irrespective of number of invitations.

Secondary outcomes measuring safety and efficacy reported here include STDR, visual acuity (VA, recorded as log of the minimum angle of resolution), visual impairment (VA ≥ +0.30 and ≥0.50), screen positive and rates of retinopathy treatment. Quality-adjusted life years (QALYs) was used to produce cost effectiveness estimates.

Statistical analysis

Our primary hypothesis was that the attendance rates at first follow-up in the two arms were equivalent with a 5% equivalence margin. The estimated minimum sample size was 4460 (90% power, 2.5% one-sided type 1 error, assuming the same attendance rate in both arms and allowing for 6%/annum loss over 24 months) (further detail in ESM). Our secondary hypothesis was that STDR detection was non-inferior in the individualised arm at a pre-specified margin of 1.5%.

Primary equivalence and non-inferiority analyses followed a per protocol (PP) approach (see ESM Results for intention-to-treat (ITT) analyses).[18] Adherence to protocol for attendance was considered at the first follow-up visit and by 24 months (+90 days) for

STDR. Multiple imputations, generated using generalised linear models (GLM) dependent on the baseline characteristics (PROC MI, SAS), assessed the effect of missing values on both PP and ITT datasets.

Within the three risk groups of the individualised arm, equivalence in attendance rates between the two arms and non-inferiority in detection of STDR were explored. Participants in the control arm were allocated to risk groups based on the RCE risks at baseline. GLMs were fitted with arm, level of risk and their interaction added as factors.

Health economics

The costs of routine screening were measured using a mixed micro-costing and observational health economics analysis over a 2-year time horizon. Societal costs, including participant and companion costs, collected using a bespoke questionnaire, comprised time lost from work (productivity losses) and travel and parking costs. A detailed workplace analysis, measuring resources and staff time to deliver the screening programme, was observed at each screening centre. This ingredient-based bottom up approach enabled a current resource-based cost to be attributed to the individual patient cost of screening, taking into account both attendees and the related cost of non-attendance. We estimated the additional costs of running the RCE using a screen population of size 22,000 (Liverpool). Treatment costs were excluded as the 2-year time horizon was felt to limit any inference that could be attributed to life time cost.

A sample of the first participants enrolled into the RCT completed EQ-5D-5L [19] and Health Utilities Index Mark 3 (HUI3) [20] questionnaires at baseline and follow-up visits. Health state utilities were mapped [21] from the EQ-5D-5L to the 3L and used a UK population tariff.[22] We applied a relevant Canadian tariff [23] to health state classifications of the HUI3 in absence of an English or UK valuation set. Discounting was not applied as both costs and QALYs were assumed to be assigned and incurred on an annual basis.

We conducted multiple imputation (MI) of chained equations using available case data and followed guidance for best practice.[24] QALYs were derived using AUC, and incremental effects estimated through OLS regression (for the univariate distributions of complete cases) and seemingly unrelated regressions (for the joint distributions of multiply imputed sets) on baseline utilities. We present unadjusted estimates as sensitivity analyses. We bootstrapped these regressions to characterise sampling distributions and derive 95% bias-corrected CIs around trial arm means and mean differences.[25] ITT analyses were conducted in STATA 16SE from an NHS/societal perspective, and post-MI analyses followed Rubin's combination rules for estimation within multiply imputed sets.[26] Further detail is available in the ESM.

Results

Figure 1 summarises the trial profile showing the numbers for eligibility, allocation, withdrawals, PP and ITT datasets for the primary analysis. From May 1, 2014, 4538 participants were enrolled; 4 withdrew after randomisation requesting removal of their trial data. Reasons for non-consent are shown in Supplementary Table 1. Allocations were 2269 to the control and 2265 to the individualised arms (198, 211, and 1856 to high medium and low risk) respectively. Last follow-up was on September 5, 2018.

Baseline characteristics of participants in the PP dataset are given in Table 1 (similar distributions for ITT in Supplementary Table 2). Participants were aged between 14 and 100 (median 63) years, 60.4% were male, 94.6% white, and 88.5% had type 2 diabetes. Those in the high-risk group were more likely to have type 1 diabetes, diagnosed with diabetes for longer, higher HbA1c and less likely to have ever smoked. Proportions with any retinopathy by group within the individualised arm were: 6 months 99.5%, 12 months 79.1%, 24 months 3.9%.

182 (4.0%) participants withdrew from the trial before first follow-up, 25 (0.6%) withdrew consent, 15 (0.3%) discontinued the intervention and 142 (3.2%) were lost to follow-up (Supplementary Table 3). Withdrawals of consent were higher in the individualised arm (0.9% vs 0.2%). Loss to follow up was higher in the individualised arm (101 (4.5%) vs 41

(1.8%)), largely due to the longer follow-up period of 24 months in the low-risk group (83.4% of the individualised arm). 15 participants prematurely discontinued the intervention, all in the individualised arm and mostly in the 24 month group.

Attendance rates at first follow-up for the control and individualised arms were 84.7% (1883/2224) and 83.6% (1754/2097) respectively (difference in proportions -1.0, 95% CI -3.2 to 1.2, PP analysis). Against the predefined acceptability margin of 5% these were regarded as equivalent (Figure 2 upper panel and Table 2). Protocol deviations excluded from this analysis were observed in 31 participants in the individualised arm; no safety effect occurred (1 participant was assigned to 12 instead of 6 months and 30 were assigned to 6 or 12 months instead of 24 months). Similar results were obtained from the ITT. PP and ITT analyses with multiple (Table 2) and simple (Supplementary Table 4) imputation confirmed equivalence in attendance rates between the two arms.

Figure 2 (lower panel, data in Table 2) shows the equivalence analysis within the individualised arm. Equivalence in attendance rates at the first follow-up visit was found for the low-risk group (control 85.7%, individualised 85.1%, difference -0.6% (-2.9 to 1.7)). For the medium-risk group, the difference in attendance rates was also very small (control 81.7%, individualised 82.2%, difference 0.6% (-7.3 to 8.4)); equivalence was not confirmed due to the relatively wide confidence interval. Attendance rates were lower in the high-risk group (control 77.3%, individualised 72.3%, difference 5.0% (-13.6 to 3.5)). Attendance rates were lower in the high-risk group (control 77.3%, individualised 72.3%, difference 5.0% (-13.6 to 3.5)) and equivalence was not observed. The attendance rates observed over 12 months however (≥1 attended appointment) were higher in the individualised arm (89.1%) compared to control (77.3%). A post hoc analysis of attendance over 24 months gave similar results (Supplementary Table 5).

The mean number of appointments per person by baseline risk allocation over 24 months was: high 1.83, medium 1.06, low 0.85. At least one change in allocation from baseline was recorded as follows: high-risk 48/160 (30.0%) to a longer interval; medium-risk 34/200 (17.0%) to shorter and 84/200 (42.0%) to a longer interval; low-risk 142/1694 (8.4%) to a shorter interval (Supplementary Table 6).

There was no evidence of a loss of ability to detect STDR over 24 months from baseline in the individualised arm (28/1956; 1.4%) compared to control (35/2042; 1.7%) with a difference -0.3 (-1.1 to 0.5) (Table 2). Non-inferiority was found for the low-risk group (control 0.6%, individualised 0.2%, difference -0.3, (-0.9 to 0.1)). For the high- and medium-risk groups non-inferiority was not confirmed, likely due to small numbers. Similar results were obtained with ITT and multiple and simple imputation.

Four participants required treatment within 6 months of being screen positive, two for STDR (one control, one high risk group), and two for non-DR reasons.

Withdrawals, premature discontinuation and loss to follow-up within 24 months showed a similar distribution across arms (Supplementary Table 8).

Further safety data are presented in the ESM and Supplementary Table 9. No effect on visual function or glycaemic control was observed.

We investigated the efficacy of individualised screening using data on numbers of attended appointments and rates of screen positive events across 24 months (Supplementary Table 9). In the individualised arm 43.2% fewer screening attendances were required (2008 vs 3536). Higher rates of screen positive by screen episode attended were seen in the individualised arm (control 4.52% (160/3536), individualised 5.08% (102/2008)). Within the individualised arm the high-risk group had the highest screen positive rate (high 10.72% (34/317), medium 6.02% (15/249), low 3.7% (53/1442)). In the high-risk group most of the screen positive cases were due to other eye disease; the rates of screen positive for DR were low at 0.5% (7/1442)). Screening episodes that detected STDR were earlier in the individualised compared to the control arm: 6-12 months 17.9% vs 2.9%; 12-18 months 32.1% vs 60.0%.

868 participants completed the questionnaires. Supplementary Table 11 presents the summary costs (2019/20 values) associated with the screening programme: NHS cost per attendance £28.73, per non-attendance £12.73; additional productivity losses and OOP payments by the patient £9.00. Summary health economic and cost effectiveness data over

the 2-year time horizon are reported in Table 3, and additional data in Supplementary Tables 12 and 13. MI supported the strict dominance of individualised screening in QALY gain and cost savings. Here we briefly summarise the results reporting conservative data from analysis of complete case QALYs and MI costs. Mean incremental QALY scores did not show a statistically significant difference between trial arms EQ-5D 0.006 (95% CI -0.039 to 0.06), EQ-VAS 0.004 (-0.049 to 0.052), and HUI3 -0.017 (-0.083 to 0.04), with agreement between societal preferences (EQ-5D/HUI3) and individual preferences (EQ-VAS). Incremental cost savings per participant were: NHS perspective £17.34 (17.02 to 17.67); societal perspective £23.11(22.73 to 23.53). The individualised arm showed incremental savings across all domains. The NHS perspective cost effectiveness plane (Figure 3) for the EQ-5D and HUI3 shows the dominance of the intervention arm in cost savings and expected maintenance of quality of life. Whilst the intention had been to report cost effectiveness acceptability curves, the dominance in cost reduction of risk-based screening and little fluctuation in QALYs across all instruments rendered this metric uninformative as the proportion cost effective was inelastic to varying thresholds. Further details of the results are available in the ESM health economics results and discussion.

Discussion

Our study shows that individualised risk-based variable-interval screening appears to be safe. Attendance at first follow-up visit was equivalent and secondary safety findings on detection of STDR, visual function and glycaemic control are supportive. Our approach reduced the number of appointments by over 40%. There were no detectable effects on QOL measures and convincing cost savings.

Important strengths of our study are the RCT design, its size, independent oversight, and direction from an expert PPI group. We utilised the emerging technology of risk calculation, based on integrating local clinical data in estimating risk, to introduce a personalised approach. We implemented a mixed Markov RCE into a clinical trial setting. Protocol deviations were few (1.4% for primary outcome) and only moderate numbers of withdrawals and loss to follow up, inevitable in a large RCT such as this. Findings were similar in our PP, ITT and MI analyses, which followed current guidance on equivalence studies.[18]

Generalisation of our findings has some limitations. Ours is a single programme that has been running for over 30 years consequently with relatively low rates of baseline DR and progression to STDR. Participants' glycaemia and BP control were relatively good and might bias the sample. Low rates of DR have been reported in other settings like ours [27,28] but our results should be treated with caution in areas with higher prevalence, poorer control of diabetes, wider ethnic group representation, or in programmes in set-up.

A move to longer intervals for people at low risk has been suggested [29-31] but without convincing evidence on safety.[10] 22% of people invited to take part in our study explicitly stated that they wished to remain in annual screening or did not want a change of interval (Supplementary Table 1). Health professionals fear that extending screen intervals may reduce perceptions of the importance of screening, leading to loss of engagement and worse diabetes care. We did not detect a worsening in glycaemic control. Our findings give substantial reassurance that a 24-month interval for a low-risk PWD in a setting such as ours is safe. However, for resource poor or rural settings in low- and middle-income countries further research is required before longer intervals can be contemplated.

Using a risk-based approach allows personalisation, offering better targeting of high risk groups and improved patient engagement. Around 15% of people had at least one change in risk-based interval: 59% people allocated to attend annual (current standard) experienced a change of interval. Aspelund and colleagues [32] have developed a similar risk engine for DR screening using risk factor data from the 1990s and conducted external validation in a Dutch cohort.[33] The strength of our RCE is that it was populated with local data and can be regularly updated with current data to reflect changing local progression rates.

For our approach to be more widely adopted assessment in other local and national screening programmes will be required. Some systems development from our research setting to an implementation environment will be required. Evaluation should include the

effect of factors such as the unexplained heterogeneity between screening programmes seen in England in grading outcomes and screening uptake.

We targeted high risk people through a 6-month interval. Attendance rates in this group were lower in the individualised arm (72.3%) compared to controls (77.3%), but the shorter interval allowed more frequent screening and earlier detection of disease. There were higher relative rates of STDR detection in the high and medium-risk individualised groups (13.4% and 3.9%) compared to 1.7% in controls with very low screen positive DR rates in the low-risk group (0.5%). Our study was powered for equivalence with all risk groups combined and not for the risk-group comparisons. Despite the hypothesis of equivalence not being supported in the high-risk group, the attendance rates observed over a period of 12 months were considerably higher in the individualised arm compared to control (89.1% versus 77.3%).

The value of adding in systemic risk factor data has been a matter of debate. It has proved difficult in several settings to reliably link primary and secondary care data due to issues with data ownership and IT system management. We overcame this through strong support from local health commissioners and primary care research groups. We needed to develop bespoke data processing, imputation and data validation. We recognise that this adds cost to screening programmes and we included this in our cost effectiveness analysis. We believe that the advantages for patient and clinician engagement outweigh the added cost and complexity.

In the UK, stratified screening based solely on retinopathy levels has been suggested.[9] Our data suggest that retinopathy levels, although a strong driver, are not the only important factors. A post hoc analysis of our RCT dataset estimated that this stratification approach[9] would allocate 66.9% of cases at baseline to a 24-month interval (82.3% in ISDR) and 33.1% to 12 months, suggesting potentially better cost effectiveness. Introduction of clinical risk factors will reinforce the message to people with diabetes that their control is crucial to the management of sight and life threatening complications of diabetes. Our PPI group advised strongly during the design phase that including clinical data was important to them.

Our data show convincing evidence that an individualised approach provides considerable cost savings compared to annual screening. By moving to risk-based variable-interval screening, patients were not compromised on quality of life. Incremental screening cost savings of £17.34 were achieved (NHS perspective), rising to £23.11 (societal perspective) per participant over the 2 years. In a screening population such as in Liverpool (22,909 invitations in 2018/19), this may amount to annual savings in the region of £199,000. For England (screening population 2.76 million (2018-19)[34]) this could amount to around £23.9m in the NHS, rising to £31.9m from a societal perspective. Such resources could be used to target groups that are hard to reach and those at high risk of visual impairment, and more cost efficiently screen the expanding population of PWD. Patients in low risk groups would be spared the inconvenience and additional personal cost of attending superfluous appointments.

The large number of observations and accuracy of the true resource cost of screening is a strength of our cost effectiveness analysis. The work could have been further strengthened by taking a long-term time horizon and including costs of treatment and blindness averted. Collection of QoL data on every patient in the study would have strengthened the analysis; our sample size was chosen to minimise disruption in the screening clinic. While methods of multiple imputation involve varied assumptions, the agreement between our complete case and multiply imputed QoL data, across instruments and adjustments, is encouraging in viewing individualised screening as a cost minimiser.

Our data on efficacy shows a higher efficiency of the individualised approach with a greater proportion of screening episodes being positive (5.1% vs 4.5%). A number of benefits include a lower burden of appointments, earlier detection of STDR for high risk people and increased capacity to see people when they are newly diagnosed. Further, in the era of personalised care a shortened interval in high risk people may increase focus on risk factor control and improve engagement with screening.

For people identified by our RCE as at low risk the rates of screen positive for DR were very low at 0.5% and even lower for STDR at under 0.2%. This is likely to apply to other

established screening programmes but should not be applied in territories without established systematic screening where the first pass prevalence will be high. The design of our study and the concerns around safety restricted us to a 24 month maximum screening interval. However our data suggest that extending intervals beyond 2 years would be reasonable, subject to further monitoring on attendance.

In conclusion, our study, the largest RCT performed to date in ophthalmology or screening, shows that all parties involved in diabetes care can be reassured that extended and personalised interval screening can be safely and effectively introduced in established systematic screening programmes. It is also applicable to other settings where clinical data is available such as in healthcare delivery organisations and integrated ophthalmology and endocrinology teams. Where current recommendations are for annual screening we provide strong evidence to support a move to variable intervals with substantial reductions in cost.

Acknowledgements: We would like to acknowledge the outstanding attention to detail from the ISDR administrative team, the CTRC, the North West Coast Clinical Research Network and our other trained researchers for their immeasurable help in recruiting to target and the LDESP team for their support and assistance. We would particularly like to thank our PPI members for their invaluable input and enthusiastic support for the trial.

Funding: This study was funded by the UK National Institute for Health Research (NIHR) [Programme Grants for Applied Research (RP-PG-1210-12016)]. The views expressed are those of the authors, not necessarily those of the NIHR or the Department of Health and Social Care. The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and has final responsibility for the decision to submit.

Role of the study sponsor: The study was sponsored by The Royal Liverpool and Broadgreen University Hospitals NHS Trust who approved the study design and protocol and provided membership of the Project Steering Committee.

Authors' relationships and activities: All authors declare grant support from the National Institute for Health Research (NIHR) UK for the submitted work and no other relationships or activities that could appear to have influenced the submitted work. Professor Mark Gabbay is part-funded by the NIHR Collaboration for Leadership in Applied Health Research and Care North West Coast.

Contributors

DMB (trial PI), MJ, IMS, JR, JPV, MG, MGF and SPH (programme CI) obtained funding and designed the trial. DMB and AW led recruitment and managed all aspects of the trial. TM led the trial management for the CTRC. MGF led the statistical team. CPC and MGF conducted statistical analysis and produced the tables apart from those in the health economics analyses, with input from IMS. MJ and JGL conducted the cost effectiveness analysis and wrote the relevant sections and tables/figures. JR led the involvement of the PPI group. SPH and DMB drafted and revised the manuscript. All members of the writing committee read and approved the final manuscript. ISDR Study Group collaborators reviewed the final manuscript. DMB is guarantor for the submission.

Data sharing statement

A fully anonymised dataset with supporting data dictionary will be available from the corresponding author 3 months after publication date for 3 years to recognised research institutions subject to approval by the ISDR Data Governance Committee of an analysis plan, a data access agreement, appropriate acknowledgment, and funding for additional costs. Results will be disseminated to patient organisations.

ISDR Study Group

Authors: Deborah M Broadbent (Trial PI), Christopher P Cheyne, Mark Gabbay, Marta García-Fiñana, Simon P Harding (Study Group Chair), Marilyn James, James G Lathe, Tracy Moitt, John R Roberts, Irene M Stratton, Jiten P Vora, Amu Wang. Collaborators: ISDR investigators: Anthony C Fisher, Paula Byrne, Daniel Seddon, Paula Williamson: ISDR Research staff: Duncan Appelbe, Ayesh Alshukri, Antonio Eleuteri, Christopher Grierson, Lola Howard, Susan U Howlin, Mehrdad Mobayen-Rahni, Andy Ovens, Christopher J Sampson, Kate Silvera, David Szmyt, Clare Thetford, Pilar Vazquez-Arango, Abigail E Williams; Patient and Public Involvement Group: John Collins, John Kelly, Peter Lees, Sandra Lees, Emily Doncaster, Betty Williams; Programme and Independent Data Safety Committees: Helen Cooper, Gideon Smith, Vineeth Kumar, Chris Rogers, Alison Rowlands, Julia West, Naveed Younis, Nathalie Massat, Catey Bunce; Liverpool Diabetic Eye Screening Programme: Ticiana Criddle, Stephanie Perrett: Liverpool Clinical Commissioning Group: Lisa Jones.

References

1. The Saint Vincent Declaration on diabetes care and research in Europe. *Acta Diabetologia* 1989;10(suppl):143-144.
2. Flaxel CJ, Adelman RA, Bailey ST, Fawzi A, Lim JI, Vemulakonda GA, Ying G-s, Diabetic retinopathy preferred practice pattern. *Ophthalmology* 2020;127(1):66-145.
3. Hammes HP, Lemmen KD, Bertram B. Diabetic retinopathy and maculopathy. *Exp Clin Endocrinol Diabetes* 2014;122:387-390
4. International Diabetes Federation Diabetes Atlas. Ninth edition. 2019 www.diabetesatlas.org
5. Younis N, Broadbent DM, Vora JP, Harding SP. Incidence of sight threatening retinopathy in patients with type 2 diabetes in the Liverpool Diabetic Eye Study: a cohort study. *Lancet* 2003;361:195-200.
6. Looker HC, Nyangoma SO, Cromie DT, et al. Predicted impact of extending the screening interval for diabetic retinopathy: the Scottish Diabetic Retinopathy Screening programme. *Diabetologia* 2013;56:1716–25.
7. Grauslund J, Andersen N, Andresen J, et al. Evidence-based Danish guidelines for screening of diabetic retinopathy. *Acta Ophthalmol* 2018:96:763–769.
8. Agardh E, Tabayat-Khani P. Adopting 3-year screening intervals for sight-threatening retinal vascular lesions in Type 2 diabetic subjects without retinopathy. *Diabetes Care* 2011;34:1-2.
9. Stratton IM, Aldington SJ, Taylor DJ, Adler AI, Scanlon PH. A simple risk stratification for time to development of sight-threatening diabetic retinopathy. *Diabetes Care* 2013;36:580-5.
10. Taylor-Phillips S, Mistry H, Leslie R et al. Extending the diabetic retinopathy screening interval beyond 1 year: a systematic review. *Br J Ophthalmol* 2016;100:105-114.
11. National Health Executive. Backlog of 150,000 cervical screenings revealed as all major health screenings failing to hit targets. Feb 01, 2019. http://www.nationalhealthexecutive.com/News/backlog-of-150000-cervical-screenings-revealed-as-all-major-health-screenings-failing-to-hit-targets-/220549 (accessed May 09, 2019)
12. Justinia T. The UK's National Programme for IT: why was it dismantled? *Health Serv Manage Res* 2017:30:2-9.
13. Klonoff DC. Personalized medicine for diabetes. *J Diabetes Sci Technol* 2008:2:335-341.
14. Muir Gray JA. The shift to personalised and population medicine. *Lancet* 2013;382:200-201
15. Broadbent DM, Sampson CJ, Wang A, et al. Individualised screening for diabetic retinopathy: the ISDR study -rationale, design and methodology for a randomised controlled trial comparing annual and individualised risk-based variable-interval. *BMJ Open* 2019;**0**:e025788. doi:10.1136/bmjopen-2018-025788
16. Individualised Screening for Diabetic Retinopathy (ISDR). Institute of Ageing and Chronic Disease. University of Liverpool. https://www.liverpool.ac.uk/ageing-and-chronic-disease/research-groups/clin-studies-eye-health/projects/individualisedscreeningfordiabeticretinopathyisdr/

17. Eleuteri A, Fisher AC, Broadbent DM, et al. Individualised variable interval risk-based screening for sight threatening diabetic retinopathy – the Liverpool Risk Calculation Engine. *Diabetologia* 2017;60:2174–2182.
18. Piaggio G, Elbourne DR, Pocock SJ, Evans SJW, Altman DG, for the CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012;308:2594-2604.
19. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L*). Qual Life Res* 2011;20(10):1727–36.
20. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics* 1995;7:490–502.
21. Hernandez-Alava, M., and S. Pudney. "eq5dmap: A command for mapping between EQ-5D-3L and EQ-5D-5L". *Stata Journal* 2018;18:395-415.
22. Dolan, P. Modelling valuations for EuroQol health states. *Medical Care* 1997;35:1095–1108.
23. Furlong W, Feeny D, Torrance G, et al. Multiplicative multi-attribute utility function for the Health Utilities Index Mark 3 (HUI3) system: a technical report. Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, Canada; 1998.
24. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:37.
25. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation. *Stat Med* 2018;37:2252–2266.https://doi.org/10.1002/sim.7654
26. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc*. 1996;91:473–489.
27. Echouffo–Tcheugui JB, Ali MK, Roglic G, Hayward RA, Narayan KM. Screening intervals for diabetic retinopathy and incidence of visual loss: a systematic review. *Diabet Med* 2013;30:1272–1292.
28. Jones S, Edwards R. Diabetic retinopathy screening: a systematic review of the economic evidence. *Diabet Med* 2009;27:249–56.
29. Olafsdottir E, Stefansson E. Biennial eye screening in patients with diabetes without retinopathy: 10-year experience. *Br J Ophthalmol* 2007;91:1599–1601.
30. Basu S, Sussman JB, Berkowitz SA, et al. Validation of risk equations for complications of type 2 diabetes (RECODe) using individual participant data from diverse longitudinal cohorts in the US. *Diabetes Care* 2018;41:586-595.
31. Leese GP, Stratton IM, Land M, Bachmann MO, Jones C, Scanlon P et al. Progression of diabetes retinal status within community screening programs and potential implications for screening intervals. *Diabetes Care* 2015;38:488-494.
32. Aspelund T, Thornórisdóttir O, Olafsdottir E, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. *Diabetologia* 2011; 54: 2525-32.
33. van der Heijden AA, Walraven I, et al. Validation of a model to estimate personalised screening frequency to monitor diabetic retinopathy. *Diabetologia* 2014;57:1332-8.
34. Public Health England. Screening KPI data: Q2 2018- 2019. https://www.gov.uk/government/publications/nhs-screening-programmes-kpi-reports-2018-to-2019 (accessed May 09, 2019)

Legends for figures

Figure 1. Consort 2010 flow diagram

Figure 2. Difference in the proportion of participants attending the first follow-up visit between the two arms. Point estimates and 90% confidence limits are provided. Upper panel: primary analysis. Lower panel: high, medium and low risk groups within the individualised arm

Figure 3.

Supplementary Figure 1. Control arm distributions (frequency) of scheduled screening appointments

Supplementary Figure 2. Individualised arm distributions (frequency) of scheduled screening appointments

Legends for tables

Table 1. Participant baseline characteristics by arm and screening interval allocation 4503 participants in the per protocol dataset

Table 2. Results of the test for equivalence in attendance rate at 1st follow-up visit and of the non-inferiority in STDR detection within 24 months, based on the per-protocol, intention-to-treat and multiple imputation datasets.

Table 3. Within-trial intention-to-treat QALYs and costs: individualised vs. annual screening

Supplementary Table 1. Reasons for consent not being provided

Supplementary Table 2. Participant baseline characteristics by arm and screening interval allocation in 4534 participants in the ITT dataset

Supplementary Table 3. Withdrawals of consent, premature discontinuation of intervention and lost-to-follow-up by arm prior to first follow-up visit (primary analysis; attendance)

Supplementary Table 4. Tests for equivalence in attendance rate at 1st follow-up visit and for non-inferiority in STDR detection within 24 months, based on simple imputation of the per-protocol and intention-to-treat datasets

Supplementary Table 5. Tests for equivalence in attendance rate at 1st follow-up visit and for non-inferiority in STDR detection within 24 months, based on simple imputation of the per-protocol and intention-to-treat datasets

Supplementary Table 6. Test for equivalence in attendance proportion over 24 months based on the per-protocol dataset (post hoc analysis).

Supplementary Table 7. Breakdown of the direction (increase/decrease) of intervals at first switch from baseline allocation over 24m by risk group in the individualised arm

Supplementary Table 8. Withdrawals of consent, premature discontinuation of intervention and lost-to-follow-up by arm within 24 months (secondary analysis; STDR detection)

Supplementary Table 9. Descriptive table summarising secondary outcomes by arm

Supplementary Table 10. Numbers of attended follow-up appointments and screen positive events within 24 months (+90 day window) by arm using the per-protocol dataset at 24 months

Supplementary Table 11. Micro-costing of the Liverpool Diabetes Eye Screening Programme

Supplementary Table 12. Control arm summary statistics of intention-to-treat health economic data at 24 months Supplementary

Supplementary Table 13. Individualised arm summary statistics of intention-to-treat health economic data at 24 months

Table 1. Participant baseline characteristics by arm and screening interval allocation in 4503 participants in the per protocol dataset

Baseline Characteristic	Arm		Baseline Risk Group**			Overall Total
	Fixed (12m)	Individualised	High	Medium	Low	
n*	2269	2234	197	211	1826	4503
Gender, n (%)						
Male	1358 (59·9)	1360 (60·9)	124 (62·9)	135 (64·0)	1101 (60·3)	2718 (60·4)
Female	911 (40·1)	874 (39·1)	73 (37·1)	76 (36·0)	725 (39·7)	1785 (39·6)
Ethnicity, n (%)						
White	2140 (94·3)	2120 (94·9)	180 (91·4)	204 (96·7)	1736 (95·1)	4260 (94·6)
Asian	48 (2·1)	30 (1·3)	2 (1·0)	3 (1·4)	25 (1·4)	78 (1·7)
Black	40 (1·8)	43 (1·9)	6 (3·0)	3 (1·4)	34 (1·9)	83 (1·8)
Chinese	7 (0·3)	6 (0·3)	1 (0·5)	1 (0·5)	4 (0·2)	13 (0·3)
Other	25 (1·1)	29 (1·3)	8 (4·1)	0 (0·0)	21 (1·2)	54 (1·2)
Unknown	9 (0·4)	6 (0·3)	0 (0·0)	0 (0·0)	6 (0·3)	15 (0·3)
Smoking Status, n (%)						
Smoker	419 (18·5)	364 (16·3)	26 (13·2)	39 (18·5)	299 (16·4)	783 (17·4)
Ex-smoker	877 (38·7)	899 (40·2)	69 (35·0)	76 (36·0)	754 (41·3)	1776 (39·4)
Non-smoker	965 (42·5)	967 (43·3)	102 (51·8)	96 (45·5)	769 (42·1)	1932 (42·9)
Unknown	8 (0·4)	4 (0·2)	0 (0·0)	0 (0·0)	4 (0·2)	12 (0·3)
Diabetes Type, n (%)						
Type 1	80 (3·5)	99 (4·4)	38 (19·3)	14 (6·6)	47 (2·6)	179 (4·0)
Type 2	2024 (89·2)	1962 (87·8)	140 (71·1)	180 (85·3)	1642 (89·9)	3986 (88·5)
Unknown	165 (7·3)	173 (7·7)	19 (9·6)	17 (8·1)	137 (7·5)	338 (7·5)
Age (years)						
Observed, n	2269	2234	197	211	1826	4503
Median (IQR)	63·3 (55·0-71·0)	62·8 (54·8-70·3)	58·3 (49·9-66·2)	60·9 (53·4-69·8)	63·7 (55·9-70·8)	63·1 (54·9-70·7)
Range	14·1-100·7	15·4-91·3	17·5-86·8	15·4-86·8	16·8-91·3	14·1-100·7
Disease Duration (years)						
Observed, n	2267	2231	197	209	1825	4498

Unknown, n	2	3	0	2	1	5
Median (IQR)	6·9 (4·2-10·9)	7·0 (4·2-11·2)	11·1 (7·3-16·1)	9·8 (6·3-13·7)	6·4 (4·0-10·1)	7·0 (4·2-11·0)
Range	0·6-66·4	1·0-44·7	1·2-44·7	1·1-37·2	1·0-39·1	0·6-66·4
HbA1c (mmol/mol)						
Observed, n	2269	2232	197	211	1824	4501
Unknown, n	0	2	0	0	2	2
Median (IQR)	51·0 (44·0-61·0)	52·0 (44·0-63·0)	67·0 (53·0-84·0)	58·0 (51·0-67·0)	50·0 (44·0-60·0)	51·0 (44·0-62·0)
Range	26·0-146·0	28·0-155·0	33·0-134·0	34·0-155·0	28·0-104·0	26·0-155·0
Systolic Blood Pressure (mmHg)						
Observed, n	2268	2234	197	211	1826	4502
Unknown, n	1	0	0	0	0	1
Median (IQR)	130·0 (121·0-138·0)	130·0 (122·0-138·0)	130·0 (124·0-138·0)	132·0 (124·0-140·0)	130·0 (122·0-138·0)	130·0 (122·0-138·0)
Range	84·0-213·0	90·0-204·0	93·0-175·0	95·0-204·0	90·0-200·0	84·0-213·0
Diastolic Blood Pressure (mmHg)						
Observed, n	2208	2180	193	201	1786	4388
Unknown, n	61	54	4	10	40	115
Median (IQR)	76·0 (70·0-80·0)	76·0 (70·0-80·0)	77·0 (70·0-80·0)	77·0 (70·0-80·0)	76·0 (70·0-80·0)	76·0 (70·0-80·0)
Range	46·0-140·0	46·0-130·0	54·0-105·0	57·0-130·0	46·0-110·0	46·0-140·0
Total Cholesterol (mmol/L)						
Observed, n	2258	2224	196	209	1819	4482
Unknown, n	11	10	1	2	7	21
Median (IQR)	4·0 (3·4-4·7)	4·0 (3·4-4·7)	4·0 (3·4-4·9)	4·0 (3·4-4·6)	4·0 (3·5-4·7)	4·0 (3·4-4·7)
Range	1·4-8·1	1·8-9·7	2·0-9·0	2·2-7·6	1·8-9·7	1·4-9·7
Retinopathy Level^, n (%)						
R0 R0	1857 (81·8)	1800 (80·6)	1 (0·5)	44 (20·9)	1755 (96·1)	3657 (81·2)
R1 R0	262 (11·5)	296 (13·2)	58 (29·4)	167 (79·1)	71 (3·9)	558 (12·4)
R1 R1	146 (6·4)	137 (6·1)	137 (69·5)	0 (0·0)	0 (0·0)	283 (6·3)

*Patients randomised who have not withdrawn or requested all data to be destroyed· **Differences across the three baseline risk groups were investigated with a statistically significant trend observed for diabetes type (p<0·0001; Cochran-Armitage test), retinopathy level (p<0·0001; Fisher's Exact test), and for age (p<0·0001), disease duration

(p<0.0001), HbA1c (p<0.0001) and systolic blood pressure (p=0.0101) using the Jonckheere-Terpstra test. A non-statistically significant trend across the three baseline groups was observed for gender (p=0.30) and ethnicity (white versus non-white, p=0.06) using the Cochran-Armitage test, smoking status (p=0.07; Fisher's Exact test), and for diastolic blood pressure (p=0.06) and total cholesterol (p=0·80) using the Jonckheere-Terpstra test. ^There were an additional 5 individuals with one eye who were randomised into the trial (0.1%) (4 had R0 in one eye and were randomised into the fixed arm (0.2% of those in the fixed arm) and 1 had R1 in one eye, was randomised to the individualised arm (<0.1% of those in the individualised arm) and allocated to 6m (0.5% of those in the 6m allocation)).

Table 2. Results of the test for equivalence in attendance rate at 1st follow-up visit and of the non-inferiority in STDR detection within 24 months, based on the per-protocol, intention-to-treat and multiple imputation datasets.

Outcome	Approach		Control arm			Individualised arm			Difference in proportions $\boldsymbol{D = p_I - p_C}$	95% CI of $\boldsymbol{D}$	
			n	Attended	Proportion Attended ($\boldsymbol{p_C}$)	n	Attended	Proportion Attended ($\boldsymbol{p_I}$)		Lower bound ($\boldsymbol{LB}$)	Upper bound ($\boldsymbol{UB}$)
Primary: Attendance at 1st follow-up	Per-protocol	Overall	2224	1883	0.847	2097	1754	0.836	-0.010	-0.032	0.012
		High risk	203	157	0.773	195	141	0.723	-0.050	-0.136	0.035
		Medium risk	169	138	0.817	208	171	0.822	0.006	-0.073	0.084
		Low risk	1852	1588	0.857	1694	1442	0.851	-0.006	-0.029	0.017
		Multiple imputation	2269	1910	0.842	2234	1870	0.837	-0.005	-0.026	0.017
	Intention-to-treat	Overall	2224	1883	0.847	2143	1798	0.839	-0.008	-0.029	0.014
		Multiple imputation	2269	1913	0.843	2265	1903	0.840	-0.004	-0.025	0.018
Outcome	Approach		n	STDR Detected	STDR Proportion ($\boldsymbol{p_C}$)	n	STDR Detected	STDR Proportion ($\boldsymbol{p_I}$)	$\boldsymbol{D}$	95% CI* of $\boldsymbol{D}$	
										$\boldsymbol{LB}$	$\boldsymbol{UB}$
Secondary: STDR within 24m	Per-protocol	Overall	2042	35	0.017	1956	28	0.014	-0.003	-0.011	0.005
		High risk	176	20	0.114	127	17	0.134	0.020	-0.053	0.100
		Medium risk	157	5	0.032	179	7	0.039	0.007	-0.038	0.051
		Low risk	1709	10	0.006	1650	4	0.002	-0.003	-0.009	0.001

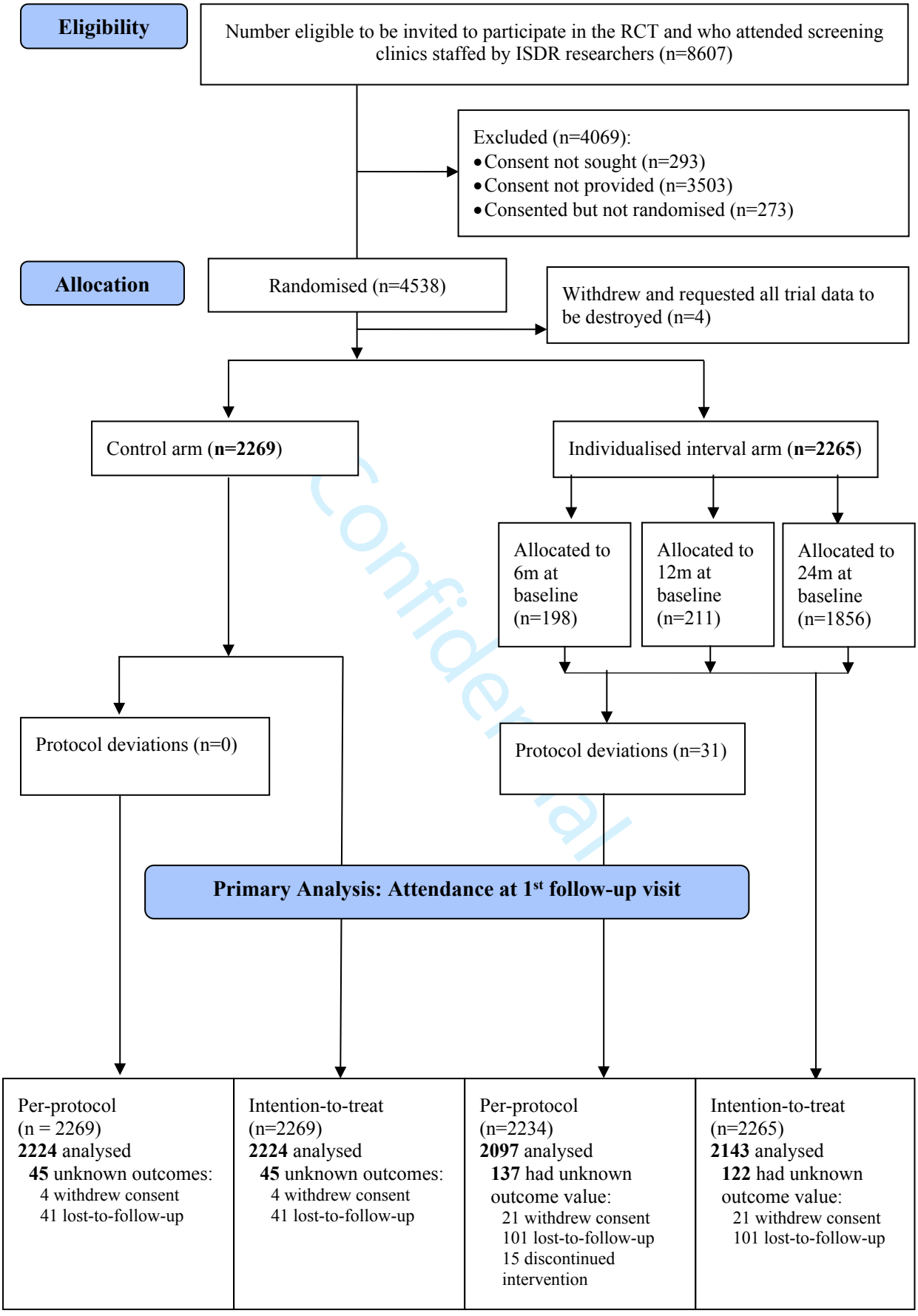
	Multiple imputation	2269	39	0.017	2234	36	0.016	-0.001	-0.009	0.007
Intention-to-treat	Overall	2042	35	0.017	2056	32	0.016	-0.002	-0.010	0.006
	Multiple imputation	2269	39	0.017	2265	39	0.017	-0.001	-0.008	0.008

*Newcombe score (based on Wilson score) confidence intervals.
The equivalence margin is predefined as δ = 0.05 and the non-inferiority margin as δ = 0.015. Proportions in both analyses (i.e., of attendance and of STDR detection) are denoted as p_C and p_I for the control and individualised arms, respectively. In both cases, results by risk group for the per-protocol approach are also included.

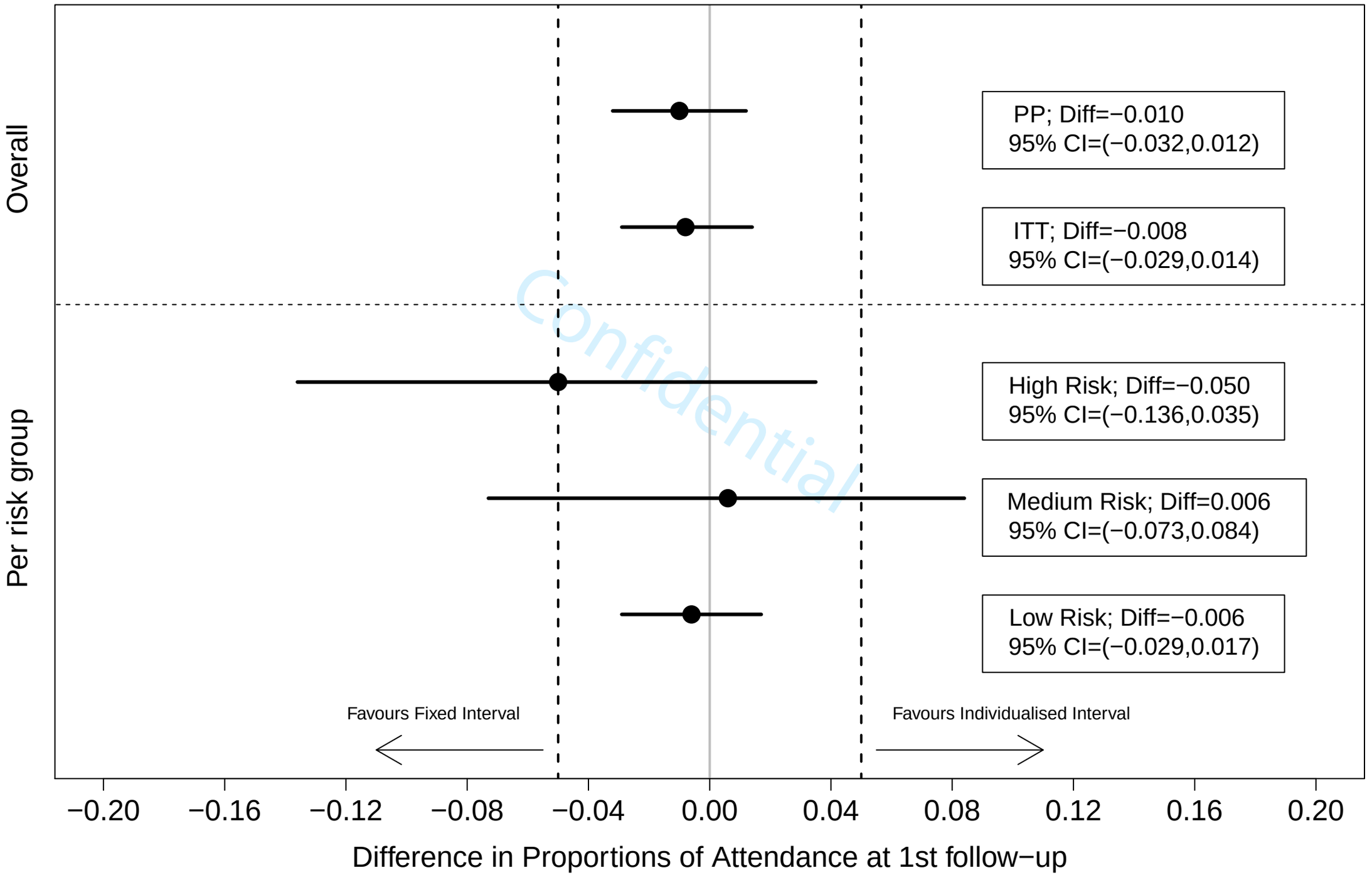
Table 3. Within-trial intention-to-treat QALYs and costs: individualised vs. annual screening

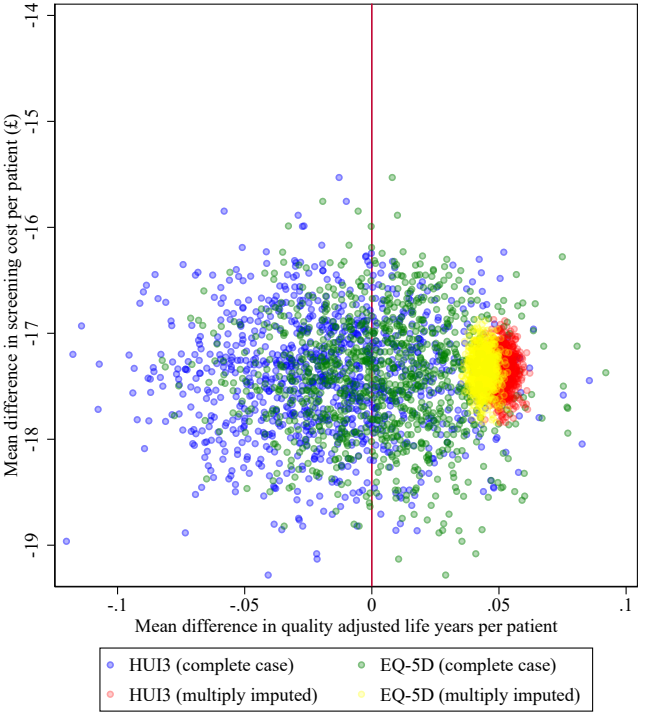
Variable (*n*/N)	Mean Difference (95% CI)	
	Complete case	Multiple imputed
EQ-5D (*539*/868)		
Unadjusted	0.012 (-0.097 to 0.119)	0.043 (0.032 to 0.055)
Baseline adjusted	0.006 (-0.039 to 0.06)	0.044 (0.038 to 0.05)
EQ-VAS (*548*/868)		
Unadjusted	-0.033 (-0.109 to 0.044)	0.013 (0.005 to 0.022)
Baseline adjusted	0.004 (-0.049 to 0.052)	0.022 (0.017 to 0.028)
HUI3 (*408*/868)		
Unadjusted	-0.016 (-0.135 to 0.116)	0.068 (0.056 to 0.081)
Baseline adjusted	-0.017 (-0.083 to 0.04)	0.051 (0.045 to 0.058)
Costs (*4389*/4534)		
NHS screening	-17.44 (-18.57 to -16.31)	-17.34 (-17.67 to -17.02)
Societal	23.26 (-24.65 to -21.92	-23.11(-23.53 to -22.73)

EQ-5D, EuroQoL-5D. EQ-VAS, EuroQol Visual Analogue Score. HUI3, Health Utilities Index Mark 3. Societal costs report the combination of NHS costs, participant or carer productivity losses, and out-of-pocket expenses. *n* corresponds to the number of univariate complete cases out of the sampled set size of N. We estimated 95% confidence intervals through 1000 iteration bootstrap regressions for univariate distributions of complete cases, and seemingly unrelated regressions for multivariate distributions of multiple imputed sets.



Equivalence Results




Mean difference in screening cost per patient (£)
-14
-15
-16
-17
-18
-19
-.1
-.05
0
.05
.1
Mean difference in quality adjusted life years per patient
HUI3 (complete case)
HUI3 (multiply imputed)
EQ-5D (complete case)
EQ-5D (multiply imputed)

Electronic Supplementary Material

Broadbent et al, Safety, efficacy and cost effectiveness of individualised screening for diabetic retinopathy: the Individualised Screening for Diabetic Retinopathy (ISDR) single centre, open label, equivalence randomised controlled trial

Table of Contents

MEMBERS OF THE ISDR STUDY GROUP

Authors: Deborah M Broadbent (Trial PI), Christopher P Cheyne, Mark Gabbay, Marta García-Fiñana, Simon P Harding (Study Group Chair), Marilyn James, James G Lathe, Tracy Moitt, John R Roberts, Irene M Stratton, Jiten P Vora, Amu Wang. Collaborators: ISDR investigators: Anthony C Fisher, Paula Byrne, John Lucy, Daniel Seddon, Paula Williamson: ISDR Research staff: Duncan Appelbe, Ayesh Alshukri, Antonio Eleuteri, Christopher Grierson, Lola Howard, Susan U Howlin, Mehrdad Mobayen-Rahni, Andy Ovens, Christopher J Sampson, Kate Silvera, David Szmyt, Clare Thetford, Pilar Vazquez-Arango, Abigail E Williams; Patient and Public Involvement Group: John Collins, John Kelly, Peter Lees, Sandra Lees, Emily Doncaster, Betty Williams; Programme and Independent Data Safety Committees: Helen Cooper, Gideon Smith, Vineeth Kumar, Chris Rogers, Alison Rowlands, Julia West, Naveed Younis, Nathalie Massat, Catey Bunce; Liverpool Diabetic Eye Screening Programme: Ticiana Criddle, Stephanie Perrett: Liverpool Clinical Commissioning Group: Lisa Jones.

ESM METHODS - TRIAL DESIGN AND CLINICAL TRIALS RESEARCH CENTRE (CTRC) PROCEDURES

Data collection and handling

The Individualised Screening for Diabetic Retinopathy (ISDR) randomised controlled trial (RCT) used a combination of data collection methods. Source data collected on paper clinical report forms (CRFs) were received centrally from the screening centres or the hospital eye service clinics by the ISDR data managers. A freeware software package designed for clinical trial data entry, Open Clinica (OC), received CRF data and systemic clinical data from the purpose built ISDR data warehouse (DW).

The ISDR trial data flowed electronically between the multiple data sources including the DW, OC, randomisation system and the risk calculation engine (RCE). Transfer of electronic data occurred within 24 hours of the scheduled visit and was considered source data for the trial unless specified differently. Allocations from the RCE were received electronically into the LDESP management system (OptoMize, EMIS Health) for generation of participant letters.

Validations were programmed into the ISDR OC system. The Statistical Team Leader, Database Designer and Data Manager or Senior Data Manager were the primary personnel involved in proposing, reviewing and amending these automated checks. These automated validations were designed to raise warnings at the point of data entry should data appear to be incorrect, out of range or invalid. Data queries were either manually raised by the OC user or automatically where data were missing or where warnings were flagged by the system.

SDV refers to source data (document) verification and normally involves checking source data against data recorded on CRFs in order to ensure quality of the trial data. For ISDR the SDV referred to comparison of data on CRFs with that entered on the database.

All (100%) of primary outcome data for randomised patients were checked against CRFs to ensure accuracy of entry.

Unknown attendance comprised withdrawal of consent, premature discontinuation of the intervention or loss to follow-up. STDR status was recorded as unknown if withdrawal occurred prior to 24 months. ITT analyses including every participant according to assigned randomisation were also conducted for sensitivity analyses.

Numbers and % described categorical variables, and mean (SD), or median (IQR) (if deviations from normality were detected) described continuous variables. Comparisons between arms for VA were conducted using the Mann Whitney U test. There was no allowance for multiplicity. All confidence intervals (CIs) were two-sided. SAS (v9.3) was used for the statistical analyses.

Trial management

A Trial Management Group met bi-monthly reporting to the Independent Data and Safety Monitoring Committee which met annually and gave safety recommendations to the Trial Steering Committee. All committees had PPI representation.

The Independent Data and Safety Management Committee (IDSMC)

IDSMC undertook annual review of the trial progress by:

- assessing data quality, including completeness
- monitoring recruitment figures and losses to follow-up
- monitoring compliance with the protocol by participants and investigators
- monitoring evidence for treatment differences in the efficacy and safety outcome measures
- recommending or advising on any major changes to the protocol, where necessary (e.g. changes to the recruitment procedures, inclusion criteria, endpoints, data collection, etc)
- requesting additional data analyses for monitoring purposes of the IDSMC
- assessing the impact and relevance of any external evidence provided

There were no stopping rules in the study, however if the IDSMC or Trial Steering Committee had any concerns with data or safety they could recommend that the study be stopped. CTRC statisticians undertook the analysis for each IDSMC report. There were no amendments other than further clarification on information and additional details in the report made by the data management committee. The Committee met on an annual basis as agreed at the first meeting and in line with recruitment and follow-up data.

Operational bias

Operational bias was minimised using the following processes:

- Automatic validations (as described above) to fire warnings at the point of data entry should data appear to be incorrect, out of range or invalid.
- Source data verification by checking the data on CRFs against that entered on the database to assure accuracy of data entry.
- A risk assessment was followed to look at safety risks from clinical procedures specified by the protocol, risks related to participant rights, risks to the reliability of trial results and organisational risks. This process was used to identify potential vulnerabilities in trial design and methodology, form a common understanding by all stakeholders on the risks for the trial, facilitate a risk proportionate approach to the trial activities, including the regulatory requirements and developing the trial management and monitoring strategies.
- The allocation sequence was generated by an independent CTRC statistician (someone different to the trial statistician and quality control (QC) statistician)
- A statistical analysis plan (SAP) was developed to describe the pre-planned analyses and reporting. SAPs were a requirement for IDSMC reports, including a closed section where allocation groups were used, and for final analyses, included reports restricted to descriptive analyses.
- A “blind review”, just before data lock, by the trial statistician where the overall (not by arm) summary statistics for the primary and secondary outcomes were assessed to detect any potential data errors
- A QC statistician (independent to the trial statistician) performed independent programming for the primary outcome and safety outcomes

Sample size calculation

Due to a slower recruitment rate than expected, the sample size required was revisited in November 2015. We considered a range of possible values for the recruitment rates based on national estimates of attendance rates, which were close to 75% and on recent figures of attendance observed at the time in the Liverpool population (minimum value just below 75%). We were aware that very small deviations from the initial 75% expected recruitment rate, for example ±0·4%, would result in differences in sample size of about ±90 patients, and consequently decided to adopt a conservative value for the expected rate to cover a range of possible scenarios. We

used Equation (9.12) [1] with allocation ratio φ=1, 10% two-sided type II error ($z_{0.95}$, power=90%), 2.5% one-sided type I error ($z_{0.975}$), equivalence margin=0.05 and equal attendance rates $\pi_1=\pi_1=\pi$=0.746. The sample size obtained, after dividing by (0.94)^2 to allow for 6% annual loss over 24 months, was n= 4460. The equivalence margin 5% was regarded by the clinical investigators as a minimum difference in attendance rates acceptable for implementation of the risk-based variable-interval approach into clinical practice and based on experience in other equivalence and inferiority trials that have resulted in changes in clinical practice.

Definition of screen positive

The full definition of screen positive in the English National Diabetic Eye Screening Programme is any of:
Moderate preproliferative diabetic retinopathy (equivalent to moderate non-proliferative) or worse, any of

- multiple deep blot haemorrhages ≥ NSC standard; venous beading; intraretinal microvascular abnormalities

Maculopathy, any of:

- exudates ≤ 1 disc diameter (DD) from the foveal centre; group exudates ≥ 1/2 disc area (DA) ≥ 1 DD from the foveal centre; haemorrhage ≤ 1 DD from the foveal centre if visual acuity ≥ +0.30 logMAR

Ungradeable images
Other significant sight threatening disease (local arrangement)

Withdrawals and missed appointments

Three levels of withdrawal were defined: withdrawal of consent (option to have all data previously recorded in the trial retained or removed), premature discontinuation of intervention (reverted to annual screening), lost to follow-up (true screen positive (STDR) and consequently transferred to the hospital eye service; moved to a different area; died; pregnancy; other). Participants who did not attend their first appointment were sent a second. If they failed to keep this they were recalled in the next round of screening (6 months for the high-risk group, 12 months for medium, low-risk and control).

Description of secondary outcomes

Due to the nature of ISDR, some outcomes refer to both safety and efficacy, and the line between the two is not easy to draw. For example detection of STDR provides evidence of safety through disease missed and efficacy through numbers of screen episodes required to detect STDR. STDR was a key patient related outcome identified by the patient expert group. We report secondary outcomes (listed on page 20 of the protocol) in Supplementary Table 9.

Three secondary endpoints: new visual impairment due to DR, number of dedicated diabetes assessment clinic appointments and the number of other eye appointments for DR, were not assessed in the final analysis as described in the SAP. These three variables refer to true screen positive patients. True screen positive patients are "withdrawn" from the trial and consequently these variables were not available to be analysed.

In addition, the following data (listed on page 31 of the protocol) proved very difficult to collect:

- number of dedicated diabetes assessment clinic appointments (biomicroscopy)
- number of other eye appointments for DR
- hospital attendance for diabetic life-threatening events (heart attack and stroke) in both arms of the trial.
- attendance at GP annual review rates in both arms of the trial.

Explanation of how the values in Supplementary Table 10 were generated

For the outcome STDR within 24 months, any individual who withdrew, prematurely discontinued the intervention or were lost-to-follow-up (except STDR cases) prior to 24 months post-baseline were considered to have an unknown STDR outcome. This is because they were not in the trial at 24 months and could have developed STDR by that time.

For the outcomes retinopathy level, maculopathy level, visual acuity and visual impairment values were taken from the last attended follow-up within 24 months (allowing for a 90-day window for attendance). In these cases, any individuals who withdrew, prematurely discontinued the intervention or were lost-to-follow-up (except those who were screen positive at the last follow-up appointment) prior to the end of the 24-month follow-up period were considered to have an unknown outcome value.

For the outcome screen positive within 24 months, any individual who withdrew, prematurely discontinued the intervention or were lost-to-follow-up (except screen positive cases) prior to 24 months post-baseline were considered to have an unknown screen positive outcome. This is because they were not in the trial at 24 months and could have been screen positive by that time.

For the outcome *false positive within 24 months*, out of those defined as being screen positive within 24 months, any individual who did not attend a biomicroscopy for any reason were considered to have an unknown false positive outcome.

ESM RESULTS - BASELINE DATA, WITHDRAWALS, PRIMARY AND SECONDARY DATA AND SUPPLEMENTARY ANALYSES

Supplementary Table 1. Reasons for consent not being provided

Status	Total	%
Consent sought	8314	100·0
Consent provided	4811	57·9
Consent not provided	3503	42·1
Does not want to take part in research	600	7·2
Does not want to be randomised	247	3·0
Does not wish to be assigned a 6 month interval	165	2·0
Does not wish to be assigned a 24 month interval	246	3·0
Patient lacking capacity	127	1·5
Already enrolled in one or more studies	18	0·2
Interpreter left due to time limitations	1	<0·1
Patient wishes to continue with annual screening	1167	14·0
No reason provided	319	3·8
Other	613	7·4

Supplementary Table 2. Participant baseline characteristics by arm and screening interval allocation in 4503 participants in the per protocol dataset

Baseline Characteristic	Arm		Baseline Risk Group**			Overall Total
	Fixed (12m)	Individualised	High	Medium	Low	
n*	2269	2234	197	211	1826	4503
Gender, n (%)						
Male	1358 (59·9)	1360 (60·9)	124 (62·9)	135 (64·0)	1101 (60·3)	2718 (60·4)
Female	911 (40·1)	874 (39·1)	73 (37·1)	76 (36·0)	725 (39·7)	1785 (39·6)
Ethnicity, n (%)						
White	2140 (94·3)	2120 (94·9)	180 (91·4)	204 (96·7)	1736 (95·1)	4260 (94·6)
Asian	48 (2·1)	30 (1·3)	2 (1·0)	3 (1·4)	25 (1·4)	78 (1·7)
Black	40 (1·8)	43 (1·9)	6 (3·0)	3 (1·4)	34 (1·9)	83 (1·8)
Chinese	7 (0·3)	6 (0·3)	1 (0·5)	1 (0·5)	4 (0·2)	13 (0·3)
Other	25 (1·1)	29 (1·3)	8 (4·1)	0 (0·0)	21 (1·2)	54 (1·2)
Unknown	9 (0·4)	6 (0·3)	0 (0·0)	0 (0·0)	6 (0·3)	15 (0·3)
Smoking Status, n (%)						
Smoker	419 (18·5)	364 (16·3)	26 (13·2)	39 (18·5)	299 (16·4)	783 (17·4)
Ex-smoker	877 (38·7)	899 (40·2)	69 (35·0)	76 (36·0)	754 (41·3)	1776 (39·4)
Non-smoker	965 (42·5)	967 (43·3)	102 (51·8)	96 (45·5)	769 (42·1)	1932 (42·9)
Unknown	8 (0·4)	4 (0·2)	0 (0·0)	0 (0·0)	4 (0·2)	12 (0·3)
Diabetes Type, n (%)						
Type 1	80 (3·5)	99 (4·4)	38 (19·3)	14 (6·6)	47 (2·6)	179 (4·0)
Type 2	2024 (89·2)	1962 (87·8)	140 (71·1)	180 (85·3)	1642 (89·9)	3986 (88·5)
Unknown	165 (7·3)	173 (7·7)	19 (9·6)	17 (8·1)	137 (7·5)	338 (7·5)
Age (years)						
Observed, n	2269	2234	197	211	1826	4503
Median (IQR)	63·3 (55·0-71·0)	62·8 (54·8-70·3)	58·3 (49·9-66·2)	60·9 (53·4-69·8)	63·7 (55·9-70·8)	63·1 (54·9-70·7)
Range	14·1-100·7	15·4-91·3	17·5-86·8	15·4-86·8	16·8-91·3	14·1-100·7
Disease Duration (years)						
Observed, n	2267	2231	197	209	1825	4498
Unknown, n	2	3	0	2	1	5

Median (IQR)	6·9 (4·2-10·9)	7·0 (4·2-11·2)	11·1 (7·3-16·1)	9·8 (6·3-13·7)	6·4 (4·0-10·1)	7·0 (4·2-11·0)
Range	0·6-66·4	1·0-44·7	1·2-44·7	1·1-37·2	1·0-39·1	0·6-66·4
HbA1c (mmol/mol)						
Observed, n	2269	2232	197	211	1824	4501
Unknown, n	0	2	0	0	2	2
Median (IQR)	51·0 (44·0-61·0)	52·0 (44·0-63·0)	67·0 (53·0-84·0)	58·0 (51·0-67·0)	50·0 (44·0-60·0)	51·0 (44·0-62·0)
Range	26·0-146·0	28·0-155·0	33·0-134·0	34·0-155·0	28·0-104·0	26·0-155·0
Systolic Blood Pressure (mmHg)						
Observed, n	2268	2234	197	211	1826	4502
Unknown, n	1	0	0	0	0	1
Median (IQR)	130·0 (121·0-138·0)	130·0 (122·0-138·0)	130·0 (124·0-138·0)	132·0 (124·0-140·0)	130·0 (122·0-138·0)	130·0 (122·0-138·0)
Range	84·0-213·0	90·0-204·0	93·0-175·0	95·0-204·0	90·0-200·0	84·0-213·0
Diastolic Blood Pressure (mmHg)						
Observed, n	2208	2180	193	201	1786	4388
Unknown, n	61	54	4	10	40	115
Median (IQR)	76·0 (70·0-80·0)	76·0 (70·0-80·0)	77·0 (70·0-80·0)	77·0 (70·0-80·0)	76·0 (70·0-80·0)	76·0 (70·0-80·0)
Range	46·0-140·0	46·0-130·0	54·0-105·0	57·0-130·0	46·0-110·0	46·0-140·0
Total Cholesterol (mmol/L)						
Observed, n	2258	2224	196	209	1819	4482
Unknown, n	11	10	1	2	7	21
Median (IQR)	4·0 (3·4-4·7)	4·0 (3·4-4·7)	4·0 (3·4-4·9)	4·0 (3·4-4·6)	4·0 (3·5-4·7)	4·0 (3·4-4·7)
Range	1·4-8·1	1·8-9·7	2·0-9·0	2·2-7·6	1·8-9·7	1·4-9·7
Retinopathy Level^, n (%)						
R0 R0	1857 (81·8)	1800 (80·6)	1 (0·5)	44 (20·9)	1755 (96·1)	3657 (81·2)
R1 R0	262 (11·5)	296 (13·2)	58 (29·4)	167 (79·1)	71 (3·9)	558 (12·4)
R1 R1	146 (6·4)	137 (6·1)	137 (69·5)	0 (0·0)	0 (0·0)	283 (6·3)

*Patients randomised who have not withdrawn or requested all data to be destroyed· **Differences across the three baseline risk groups were investigated with a statistically significant trend observed for diabetes type (p<0·0001; Cochran-Armitage test), retinopathy level (p<0·0001; Fisher's Exact test), and for age (p<0·0001), disease duration (p<0·0001), HbA1c (p<0·0001) and systolic blood pressure (p=0·0101) using the Jonckheere-Terpstra test. A non-statistically significant trend across the three baseline groups was observed for gender (p=0·30) and ethnicity (white versus non-white, p=0·06) using the Cochran-Armitage test, smoking status (p=0·07; Fisher's Exact test), and for diastolic blood pressure (p=0·06) and total cholesterol (p=0·80) using the Jonckheere-Terpstra test. ^There were an additional 5 individuals with one eye who were randomised into the trial (0·1%) (4 had R0 in one eye and were randomised into the fixed arm (0·2% of those in the fixed arm) and 1 had R1 in one eye, was randomised to the individualised arm (<0·1% of those in the individualised arm) and allocated to 6m (0·5% of those in the 6m allocation)).

Supplementary Table 3. Participant baseline characteristics by arm and screening interval allocation in 4534 participants in the ITT dataset

Baseline characteristic	Arm		Baseline risk group for individualised arm			Overall total
	Control (12m)	Individualised	High	Medium	Low	
n*	2269	2265	198	211	1856	4534
Gender, n (%)						
Male	1358 (59·9)	1375 (60·7)	124 (62·6)	135 (64·0)	1116 (60·1)	2733 (60·3)
Female	911 (40·1)	890 (39·3)	74 (37·4)	76 (36·0)	740 (39·9)	1801 (39·7)
Ethnicity, n (%)						
White	2140 (94·3)	2151 (95·0)	181 (91·4)	204 (96·7)	1766 (95·2)	4291 (94·6)
Asian	48 (2·1)	30 (1·3)	2 (1·0)	3 (1·4)	25 (1·3)	78 (1·7)
Black	40 (1·8)	43 (1·9)	6 (3·0)	3 (1·4)	34 (1·8)	83 (1·8)
Chinese	7 (0·3)	6 (0·3)	1 (0·5)	1 (0·5)	4 (0·2)	13 (0·3)
Other	25 (1·1)	29 (1·3)	8 (4·0)	0 (0·0)	21 (1·1)	54 (1·2)
Unknown	9 (0·4)	6 (0·3)	0 (0·0)	0 (0·0)	6 (0·3)	15 (0·3)
Smoking status, n (%)						
Smoker	419 (18·5)	371 (16·4)	26 (13·1)	39 (18·5)	306 (16·5)	790 (17·4)
Ex-smoker	877 (38·7)	909 (40·1)	69 (34·8)	76 (36·0)	764 (41·2)	1786 (39·4)
Non-smoker	965 (42·5)	981 (43·3)	103 (52·0)	96 (45·5)	782 (42·1)	1946 (42·9)
Unknown	8 (0·4)	4 (0·2)	0 (0·0)	0 (0·0)	4 (0·2)	12 (0·3)
Diabetes type, n (%)						
Type 1	80 (3·5)	103 (4·5)	39 (19·7)	14 (6·6)	50 (2·7)	183 (4·0)
Type 2	2024 (89·2)	1988 (87·8)	140 (70·7)	180 (85·3)	1668 (89·9)	4012 (88·5)
Unknown	165 (7·3)	174 (7·7)	19 (9·6)	17 (8·1)	138 (7·4)	339 (7·5)
Age (years)						
Observed, n	2269	2265	198	211	1856	4534
Median (IQR)	63·3 (55·0-71·0)	62·8 (54·9-70·3)	58·3 (49·6-66·2)	60·9 (53·4-69·8)	63·7 (56·0-70·8)	63·1 (54·9-70·7)
Range	14·1-100·7	15·4-95·7	17·5-86·8	15·4-86·8	16·8-95·7	14·1-100·7
Disease duration (years)						
Observed, n	2267	2262	198	209	1855	4529
Unknown, n	2	3	0	2	1	5

Median (IQR)	6·9 (4·2-10·9)	7·0 (4·3-11·2)	11·1 (7·3-16·1)	9·8 (6·3-13·7)	6·4 (4·0-10·1)	7·0 (4·2-11·0)
Range	0·6-66·4	1·0-44·7	1·2-44·7	1·1-37·2	1·0-39·1	0·6-66·4
HbA1c (mmol/mol)						
Observed, n	2269	2263	198	211	1854	4532
Unknown, n	0	2	0	0	2	2
Median (IQR)	51·0 (44·0-61·0)	52·0 (44·0-63·0)	66·5 (53·0-84·0)	58·0 (51·0-67·0)	50·0 (44·0-60·0)	51·0 (44·0-62·0)
Range	26·0-146·0	28·0-155·0	33·0-134·0	34·0-155·0	28·0-104·0	26·0-155·0
Systolic blood pressure (mmHg)						
Observed, n	2268	2265	198	211	1856	4533
Unknown, n	1	0	0	0	0	1
Median (IQR)	130·0 (121·0-138·0)	130·0 (122·0-138·0)	130·0 (124·0-138·0)	132·0 (124·0-140·0)	130·0 (122·0-138·0)	130·0 (122·0-138·0)
Range	84·0-213·0	90·0-204·0	93·0-175·0	95·0-204·0	90·0-200·0	84·0-213·0
Diastolic blood pressure (mmHg)						
Observed, n	2208	2206	194	201	1811	4414
Unknown, n	61	59	4	10	45	120
Median (IQR)	76·0 (70·0-80·0)	76·0 (70·0-80·0)	77·0 (70·0-80·0)	77·0 (70·0-80·0)	75·0 (70·0-80·0)	76·0 (70·0-80·0)
Range	46·0-140·0	46·0-130·0	54·0-105·0	57·0-130·0	46·0-110·0	46·0-140·0
Total cholesterol (mmol/L)						
Observed, n	2258	2254	197	209	1848	4512
Unknown, n	11	11	1	2	8	22
Median (IQR)	4·0 (3·4-4·7)	4·0 (3·4-4·7)	4·0 (3·4-4·9)	4·0 (3·4-4·6)	4·0 (3·5-4·7)	4·0 (3·4-4·7)
Range	1·4-8·1	1·8-9·7	2·0-9·0	2·2-7·6	1·8-9·7	1·4-9·7
Retinopathy level^, n (%)						
R0 R0	1857 (81·8)	1830 (80·8)	1 (0·5)	44 (20·9)	1785 (96·2)	3687 (81·3)
R1 R0	262 (11·5)	297 (13·1)	59 (29·8)	167 (79·1)	71 (3·8)	559 (12·3)
R1 R1	146 (6·4)	137 (6·0)	137 (69·2)	0 (0·0)	0 (0·0)	283 (6·2)

*Patients randomised who have not withdrawn and requested all data to be destroyed. ^In addition there were 5 individuals with one eye who were randomised into the trial (0·1%): 4 had R0 in one eye and were randomised into the fixed arm (0·2% of those in the fixed arm) and 1 had R1 in one eye, was randomised to the individualised arm (<0·1% of those in the individualised arm) and allocated to 6m (0·5% of those in the 6m allocation).

Withdrawals

182 (4.0%) participants withdrew from the trial before first follow-up, 25 (0.6%) withdrew consent, 15 (0.3%) discontinued the intervention and 142 (3.2%) were lost to follow-up (Supplementary Table 4). Withdrawals of consent were higher in the individualised arm (0.9% vs 0.2%). Loss to follow up was higher in the individualised arm, 101 (4.5%) compared to 41 (1.8%) in the control arm, largely due to the longer follow-up period of 24 months experienced by low risk participants (83.4% of the individualised arm). 15 participants prematurely discontinued the intervention, all in the individualised arm and mostly in the 24 month group.

Supplementary Table 4. Withdrawals of consent, premature discontinuation of intervention and lost-to-follow-up by arm prior to first follow-up visit (primary analysis; attendance)

Withdrawals, n (%)	Arm		Overall Total
	Control	Individualised	
Number of participants (per-protocol dataset*)	2269 (100·0)	2234 (100·0)	4503 (100·0)
Attendance within 90-day window known	2224 (98·0)	2097 (93·9)	4321 (96·0)
Attendance within 90-day window unknown	45 (2·0)	137 (6·1)	182 (4·0)
Withdrawal of consent	4 (0·2)	21 (0·9)	25 (0·6)
unhappy with screening interval provided, wishes to be seen more often	0 (0·0)	16 (0·7)	16 (0·4)
no longer wished to take part	2 (0·1)	4 (0·2)	6 (0·1)
did not wish to provide reason	2 (0·1)	1 (<0·1)	3 (0·1)
Prematurely discontinued the intervention	0 (0·0)	15 (0·7)	15 (0·3)
Lost-to-follow-up	41 (1·8)	101 (4·5)	142 (3·2)
moved to a non-participating GP practice	10 (0·4)	32 (1·4)	42 (0·9)
patient died	25 (1·1)	54 (2·4)	79 (1·8)
true screen positive	0 (0·0)	0 (0·0)	0 (0·0)
opted out of screening	0 (0·0)	3 (0·1)	3 (0·1)
no longer diabetic	4 (0·2)	10 (0·4)	14 (0·3)
pregnant	2 (0·1)	1 (<0·1)	3 (0·1)
other^	0 (0·0)	1 (<0·1)	1 (<0·1)

*Note that this table excludes protocol deviations (n=31).

^This patient was on an insulin pump.

Supplementary Table 5. Tests for equivalence in attendance rate at 1st follow-up visit and for non-inferiority in STDR detection within 24 months, based on simple imputation of the per-protocol and intention-to-treat datasets.
In the most conservative approach all missing values are set to "not attended" and "STDR detected" and in the least conservative approach they were set to "attended" and "STDR not detected", for the primary and secondary outcomes, respectively. The equivalence margin is predefined as δ = 0·05 and the non-inferiority margin as δ = 0·015. Proportions in both analyses (i.e., of attendance and of STDR detection) are denoted as p_C and p_I for the control and individualised arms, respectively.

Outcome	Approach		Control arm			Individualised arm			Difference in proportions $D = p_I - p_c$	95% CI of D	
			n	Attended	Proportion Attended (p_C)	n	Attended	Proportion Attended (p_I)		Lower bound (LB)	Upper bound (UB)
Primary: Attendance at 1st follow-up	**Per-protocol**	**Most Conservative**	2269	1883	0·830	2234	1754	0·785	-0·045	-0·068	0·022
		Least Conservative	2269	1928	0·850	2234	1891	0·846	-0·003	-0·024	0·018
	Intention-to-treat	**Most Conservative**	2269	1883	0·830	2265	1798	0·794	-0·036	-0·059	0·013
		Least Conservative	2269	1928	0·850	2265	1920	0·848	-0·002	-0·023	0·019
Outcome	**Approach**		**n**	**STDR Detected**	**STDR Proportion (p_C)**	**n**	**STDR Detected**	**STDR Proportion (p_I)**	**D**	**95% CI^ of D**	
										LB	***UB***
Secondary: STDR within 24m	**Per-protocol**	**Most Conservative**	2269	262	0·115	2191	263	0·120	0·005	-0·014	0·024
		Least Conservative	2269	35	0·015	2191	28	0·013	-0·003	-0·010	0·004
	Intention-to-treat	**Most Conservative**	2269	262	0·115	2265	241	0·106	-0·009	-0·027	0·009
		Least Conservative	2269	35	0·015	2265	32	0·014	-0·001	-0·009	0·006

^Newcombe score 95% confidence intervals.

Supplementary Table 6. Test for equivalence in attendance proportion over 24 months based on the per-protocol dataset (post hoc analysis)
The equivalence margin is predefined as δ = 0·05. Proportion of attendance (denoted as p_C and p_I for the control and individualised arms, respectively) was generated from a generalised linear regression model and can be interpreted as a *weighted* mean of the mean individual attendance proportion (which is calculated as the number of attended visits divided by the number of expected visits per individual).

Outcome	Approach	Control arm		Individualised arm		Difference in proportions $D = p_I - p_c$	95% CI of D	
		n	Proportion Attended (p_C)	n	Proportion Attended (p_C)		Lower bound (LB)	Upper bound (UB)
Attendance over 24 months	**Per- protocol**	2224	0·848	2054	0·843	-0·005	-0·026	0·017

This table excludes 74 protocol deviations that occurred within 24 months in the individualised arm, and 45 people in the control arm, as well as 137 people in the individualised arm for whom the attendance outcome is unknown since they withdrew consent, premature discontinued the intervention or were lost-to-follow-up prior to their 1st follow-up visit.

Supplementary Table 7. Breakdown of the direction (increase/decrease) of intervals at first switch from baseline allocation over 24m by risk group in the individualised arm

Risk group	n^	Change in interval at 1st switch from baseline allocation	At least one change in allocation (n)	Percentage of individuals with at least one change in allocation
High	160	Change to 12m or 24m	48	30.0
Medium	200	Change to 6m	34	17.0
		Change to 24m	84	42.0
Low	1694	Change to 6m or 12m	142	8.4

^Number of individuals with a recorded attendance over 24m outcome

Supplementary Table 8. Withdrawals of consent, premature discontinuation of intervention and lost-to-follow-up by arm within 24 months (secondary analysis; STDR detection)

Withdrawals, n (%)	Arm		Overall Total
	Control	Individualised	
Number of participant (per protocol dataset)*	2269 (100·0)	2191 (100·0)	4460 (100·0)
STDR within 24m known	2042 (90·0)	1956 (89·3)	3998 (89·6)
STDR within 24m unknown	227 (10·0)	235 (10·7)	462 (10·4)
Withdrawal of consent	22 (1·0)	35 (1·6)	57 (1·3)
unhappy with screening interval provided, wishes to be seen less often	0 (0·0)	1 (<0·1)	1 (<0·1)
unhappy with screening interval provided, wishes to be seen more often	0 (0·0)	26 (1·2)	26 (0·6)
burden of additional data collection	3 (0·1)	0 (0·0)	3 (0·1)
no longer wished to take part	6 (0·3)	4 (0·2)	10 (0·2)
did not wish to provide reason	13 (0·6)	4 (0·2)	17 (0·4)
Prematurely discontinued the intervention	0 (0·0)	40 (1·8)	40 (0·9)
Lost-to-follow-up	191 (8·4)	155 (7·1)	346 (7·8)
moved to a non-participating GP practice	40 (1·8)	38 (1·7)	78 (1·7)
patient died	67 (3·0)	65 (3·0)	132 (3·0)
true screen positive**	61 (2·7)	28 (1·3)	89 (2·0)
retained in hospital eye service – not STDR	4 (0·2)	4 (0·2)	8 (0·2)

opted out of screening	3 (0·1)	4 (0·2)	7 (0·2)
no longer diabetic	9 (0·4)	10 (0·5)	19 (0·4)
pregnant	4 (0·2)	2 (0·1)	6 (0·1)
Other^	3 (0·1)	4 (0·2)	7 (0·2)
Screen positive but no biomicroscopy recorded	14 (0·6)	5 (0·2)	19 (0·4)

*Note that this table excludes protocol deviations (n=74).
**Does not include STDR at follow-up as this is the outcome of interest.
^ Other reasons included: participant went into the digital surveillance pathway (n= 2 in each arm), was false positive but did not come back into the trial (n=1 in each arm) and was an insulin pump patient (n=1 in the individualised arm).

We did not detect differences in logMAR VA (p=0.64) or in rates of VI in the better eye at the last attended visit between the two arms (secondary safety outcome, Supplememntary Table 9). Findings for worse eye VA and VI ≥ +0·50 were similar.

Supplementary Table 9. Descriptive table summarising secondary outcomes by arm

Secondary outcomes	Arm		Overall
	Control	Individualised	
n	2269	2191	4460
STDR within 24m, n (%; 95% CI^)			
STDR	35 (1·5; 1·1-2·1)	28 (1·3; 0·9-1·8)	63 (1·4; 1·1-1·8)
Not STDR	2007 (88·5; 87·1-89·7)	1928 (88·0; 86·6-89·3)	3935 (88·2; 87·3-89·1)
Unknown	227 (10·0; 8·8-11·3)	235 (10·7; 9·5-12·1)	462 (10·4; 9·5-11·3)
Attendance over 24m			
In trial for at least one follow-up, n	2224	2054	4278
Did not reach first follow-up in trial, n	45	137	182
Mean of mean individual attendance, % (95% CI)	84·8 (83·4-86·3)	83·6 (82·1-85·2)	84·3 (83·2-85·3)
Retinopathy level at last attended screening appointment within 24m, n (%; 95% CI^)			
R0/R0	1509 (66·5; 64·5-68·4)	1343 (61·3; 59·2-63·3)	2852 (63·9; 62·5-65·3)
R1/R0	239 (10·5; 9·3-11·9)	177 (8·1; 7·0-9·3)	416 (9·3; 8·5-10·2)
R1/R1	156 (6·9; 5·9-8·0)	138 (6·3; 5·4-7·4)	294 (6·6; 5·9-7·4)
R2/R1	8 (0·4; 0·2-0·7)	4 (0·2; 0·1-0·5)	12 (0·3; 0·2-0·5)
R2/R2	4 (0·2; 0·1-0·5)	2 (0·1; 0·0-0·3)	6 (0·1; 0·1-0·3)
Unassessable/R0	63 (2·8; 2·2-3·5)	23 (1·0; 0·7-1·6)	86 (1·9; 1·6-2·4)

Unassessable/R1	10 (0·4; 0·2-0·8)	8 (0·4; 0·2-0·7)	18 (0·4; 0·3-0·6)
Unassessable/R2	1 (<0·1; 0·0-0·3)	0 (0·0; NA)	1 (<0·1; 0·0-0·1)
Unassessable/Unassessable	11 (0·5; 0·3-0·9)	10 (0·5; 0·3-0·8)	21 (0·5; 0·3-0·7)
Unknown	268 (11·8; 10·6-13·2)	486 (22·2; 20·5-24·0)	754 (16·9; 15·8-18·0)
Maculopathy level at last attended screening appointment within 24m, n (%; 95% CI^)			
M0/M0	1877 (82·7; 81·1-84·2)	1633 (74·5; 72·7-76·3)	3510 (78·7; 77·5-79·9)
M1/M0	35 (1·5; 1·1-2·1)	29 (1·3; 0·9-1·9)	64 (1·4; 1·1-1·8)
M1/M1	3 (0·1; 0·0-0·4)	2 (0·1; 0·0-0·3)	5 (0·1; 0·1-0·3)
Unassessable/M0	71 (3·1; 2·5-3·9)	30 (1·4; 1·0-2·0)	101 (2·3; 1·9-2·7)
Unassessable/M1	4 (0·2; 0·1-0·5)	1 (<0·1; 0·0-0·3)	5 (0·1; 0·1-0·3)
Unassessable/Unassessable	11 (0·5; 0·3-0·9)	10 (0·5; 0·3-0·8)	21 (0·5; 0·3-0·7)
Unknown	268 (11·8; 10·6-13·2)	486 (22·2; 20·5-24·0)	754 (16·9; 15·8-18·0)
Screen positives within 24m, n (%; 95% CI^)			
Screen Positive	160 (7·1; 6·1-8·2)	105 (4·8; 4·0-5·8)	265 (5·9; 5·3-6·7)
Screen Negative	1852 (81·6; 80·0-83·2)	1611 (73·5; 71·6-75·3)	3463 (77·6; 76·4-78·8)
Unknown	257 (11·3; 10·1-12·7)	475 (21·7; 20·0-23·5)	732 (16·4; 15·4-17·5)
False positives within 24m, n (%; 95% CI^)			
Screen Positive, n	160	105	265
False Positive	18 (11·3; 7·2-17·1)	14 (13·3; 8·1-21·1)	32 (12·1; 8·7-16·6)
True Positive	125 (78·1; 71·1-83·8)	81 (77·1; 68·2-84·1)	206 (77·7; 72·4-82·3)
Condition Improved	3 (1·9; 0·6-5·4)	5 (4·8; 2·1-10·7)	8 (3·0; 1·5-5·8)
Unknown	14 (8·8; 5·3-14·2)	5 (4·8; 2·1-10·7)	19 (7·2; 4·6-10·9)
Visual acuity in better eye (logMAR)			
Observed, n	2000	1706	3706
Unknown, n	269	485	754
Median (IQR)	+0·02 (+0·00, +0·10)	+0·02 (+0·00, +0·10)	+0·02 (+0·00, +0·10)
Range	[-0·30, +2·00]	[-0·30, +0·72]	[-0·30, +2·00]
Visual acuity in worse eye (logMAR)			
Observed, n	2000	1706	3706
Unknown, n	269	485	754
Median (IQR)	+0·12 (+0·02, +0·18)	+0·12 (+0·02, +0·18)	+0·12 (+0·02, +0·18)
Range	[-0·20, +3·00]	[-0·20, +3·00]	[-0·20, +3·00]

Visual impairment (VA ≥ +0·30 in better eye), n (%; 95% CI^)			
Visually Impaired	69 (3·0; 2·4-3·8)	56 (2·6; 2·0-3·3)	125 (2·8; 2·4-3·3)
Not Visually Impaired	1931 (85·1; 83·6-86·5)	1650 (75·3; 73·5-77·1)	3581 (80·3; 79·1-81·4)
Unknown	269 (11·9; 10·6-13·3)	485 (22·1; 20·5-23·9)	754 (16·9; 15·8-18·0)
Visual impairment (VA ≥ +0·50 in better eye), n (%; 95% CI^)			
Visually Impaired	16 (0·7; 0·4-1·1)	11 (0·5; 0·3-0·9)	27 (0·6; 0·4-0·9)
Not Visually Impaired	1984 (87·4; 86·0-88·7)	1695 (77·4; 75·6-79·1)	3679 (82·5; 81·4-83·6)
Unknown	269 (11·9; 10·6-13·3)	485 (22·1; 20·5-23·9)	754 (16·9; 15·8-18·0)
Number of screening appointments / Number of non-attended screening appointments, n (%; 95% CI)			
Total number of screening appointments, n	5034	3074	8108
Attended, n (%; 95% CI)	3536 (70·2; 69·0-71·5)	2008 (65·3; 63·6-67·0)	5544 (68·4; 67·4-69·4)
Did Not Attend, n (%; 95% CI)	1498 (29·8; 28·5-31·0)	1066 (34·7; 33·0-36·4)	2564 (31·6; 30·6-32·6)

^Wilson score confidence intervals.

Two participants required treatment for DR within 6 months of being screen positive (one control, one high risk group).

We did not detect a clinically significant worsening of diabetes control comparing HbA1c levels across the trial arms or between the individualised arm groups (post hoc analysis). Changes in median HbA1c (mmol/mol) from to 24 months were: control 1, individualised 2 (high 2, medium1, low 2). There was no difference in the proportions of participants in each group who had a significant (≥11mmol/mol) increase in HbA1c during the trial: control 15.2%, individualised 14.6%; high 13.4%, medium 15·0%, low 14.7%.

Supplementary Table 10. Numbers of attended follow-up appointments and screen positive events within 24 months (+90 day window) by arm using the per-protocol dataset at 24 months

Screen Positive Status	Arm		Risk Group within Individualised Arm			Total
	Control	Individualised	High	Medium	Low	
Total number of individuals, n	2269	2191	162	203	1826	4460
Number of attended follow-up appointments, n (%)	3536 (100·0)	2008 (100·0)	317 (100·0)	249 (100·0)	1442 (100·0)	5544 (100·0)
Screen positive, n (%)	160 (4·5)	102 (5·1)	34 (10·7)	15 (6·0)	53 (3·7)	262 (4·7)
Screen positive for DR, n (%)	54 (1·5)	45 (2·2)	29 (9·1)	9 (3·6)	7 (0·5)	99 (1·8)
Screen positive for unassessable images, n (%)	79 (2·2)	39 (1·9)	3 (0·9)	5 (2·0)	31 (2·1)	118 (2·1)
Screen positive for other eye disease requiring HES, n (%)	27 (0·8)	18 (0·9)	2 (0·6)	1 (0·4)	15 (1·0)	45 (0·8)
Screen negative, n (%)	3371 (95·3)	1905 (94·9)	283 (89·3)	234 (94·0)	1388 (96·3)	5276 (95·2)
Unknown status, n (%)	5 (0·1)	1 (<0·1)	0 (0·0)	0 (0·0)	1 (0·1)	6 (0·1)

ESM METHODS - HEALTH ECONOMICS

Utilities and quality-adjusted life-years

Baseline health economic data were observed for the first 868 eligible individuals, out of a total intention-to-treat study size of 4534, yet it should be emphasised this was limited data collection by design. As such these data fulfil the criteria of missing completely at random (MCAR), where the likelihood of missingness is independent of unobserved data, and will not bias complete case analysis since the reason for the missing data is also unrelated to the outcome.[2, 3] Additionally, an unforeseen software error led a number of HUI classifications to be irretrievable – the affected observations were evenly split between trial arms and can similarly be considered MCAR. Health state utilities were derived from the EQ-5D, being the measure of choice of the National Institute for Health and Care Excellence (NICE), and the HUI3 for it's inclusion of a sight specific component, as such exhibiting greater sensitivity to diabetic retinopathy attributable changes in health related QoL. Following NICE's October 2019 position statement;[4] we mapped patients' item-level health state classifications, from the EQ-5D-5L instrument, to the 3L values using an appropriate mapping tool, the EuroQoL dataset, a Copula cumulative distribution model, and UK valuation set. Our estimations for the HUI3 applied a relevant Canadian tariff in absence of an English or UK valuation set. EQ-VAS estimates were divided by one hundred to aid comparison with the other instruments. Elicited utility scores of zero were input for 39 total participants when censored due to death. These solely represent individuals who supplied baseline health economic data, and not for those unobserved data which would have introduced an association between death and the observation of health state outcomes. Of note: these patients represent the number of zero value health state utilities reported at 6 months within our summary statistics. While not all patients were normally observed at this timepoint, we chose this method to best inform QALY construction and similarly this number, and those who later died, were amended at the 12, 18, and 24 month timepoints to ensure correct granularity. Similarly, observed utilities at each appointment were assigned to their nearest 6-month interval (allowing for 90 days on either side).

Area under the curve (AUC) [5] was used to adjust participant quality of life for the time spent in their respective health states and for constructing quality adjusted life-years (QALYs). It should be noted that patients completed health state utility instruments at baseline and at each follow-up visit, so there existed a differential in the frequency of observed responses between arm allocations at each appointment i.e. patients allocated to the high risk group at baseline completed the EQ-5D-5L and HUI3 at six months, the medium risk group at 12 months, and the low risk group at 24 months. This meant we observed greater intra-patient utility variation for higher risk participants since AUC assumes a linear relationship between health state utilities at each discrete time point. For robustness we conducted parallel analysis which constructed QALYs through change from baseline; however it had little effect on our estimates as the intra-variation in the utility of high risk patients was significantly outweighed by inter-patient variation. We utilised AUC in order to include all available data points in the generation of complete case QALYs, and because it allowed the use of far greater available case data within models of multiple imputation. Discounting was not applied as both costs and outcomes (QALYs) were assumed to be assigned and incurred on an annual basis.

Patient-level baseline utility imbalances between the control and individualised arms, regardless of statistical significance or missingness, would have led to inaccurate QALY estimates.[6] To control for this confounding, a binary variable representing the personalised arm, alongside a continuous variable for baseline utility, was regressed on total QALYs to estimate mean differences, and adjusted at means in our presented summary statistics. Means, standard deviations, and visual histogram plotting were used to compare complete versus available case utility data, exhibiting no systematic differences in values or distributions and so supporting the suitability of this method.[7] We present unadjusted estimates alongside as sensitivity analysis (covariate adjustment in addition to baseline adjustment had little effect on mean differences). The summary statistics we report are univariate distributions since deriving the multivariate distributions of complete case data of differential observational set sizes would necessarily exclude a significant number of observations from our larger QALY and costing sets.

Costs

Across multiple screening sites, a mixed methods study examined the costs of screening for sight-threatening diabetic retinopathy within the Liverpool Diabetic Eye Screening Programme (LDESP). This involved identifying the processes and patient pathways associated with photographic screening, measuring resource use, estimating the average cost of resource use, and the average cost of each patient pathway. Data were retrieved from the ISDR clinical data warehouse of screening appointments attended by participants during the 2016 calendar year. Local resource use data were obtained from staff at the screening sites across the

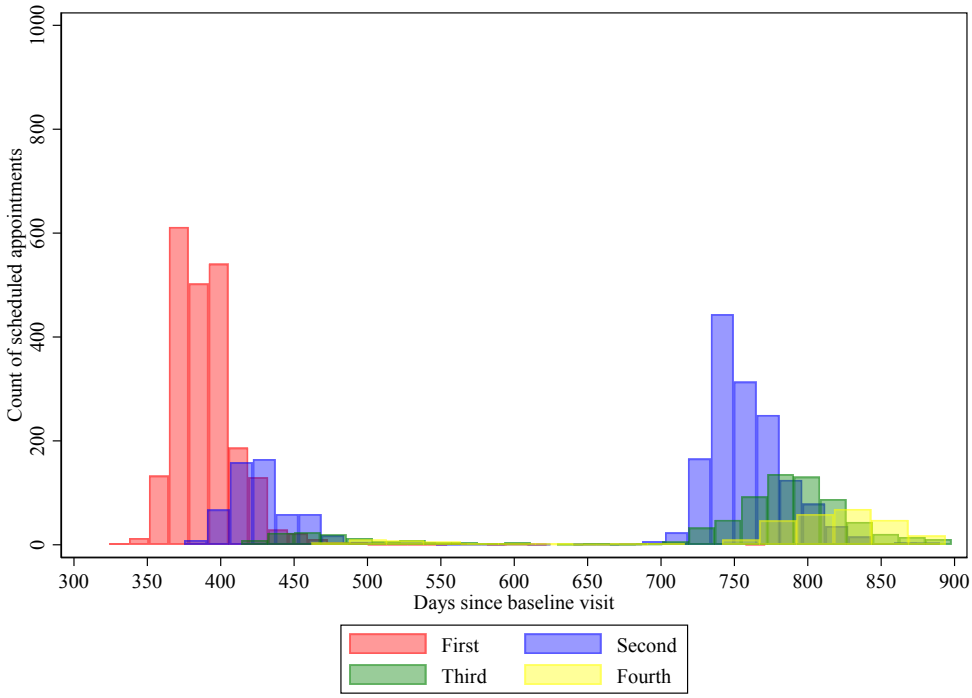
LDESP programme. The costing of resource use, such as the banding of NHS staff salaries, used unit costs derived from the PSSRU unit costs of health and social care 2017.[8] The ISDR visit questionnaire, administered to the aforementioned 868 participants for whom we observed baseline health economic data, informed the cross-sectional study of productivity losses and out-of-pocket costs to participants and carers associated with the attendance at screening.

These costs were inflated at the midpoint of the NHS financial year; an increase of 4.373% from October 2017 to October 2019 specified by the inflationary rates of the Office for National Statistics[9] for Medical Services and Paramedical Services (CDID: D7FA). Following input from the trial management group and stakeholders, we estimated the additional costs of running the RCE using a screening population of size 22,000 (Liverpool).

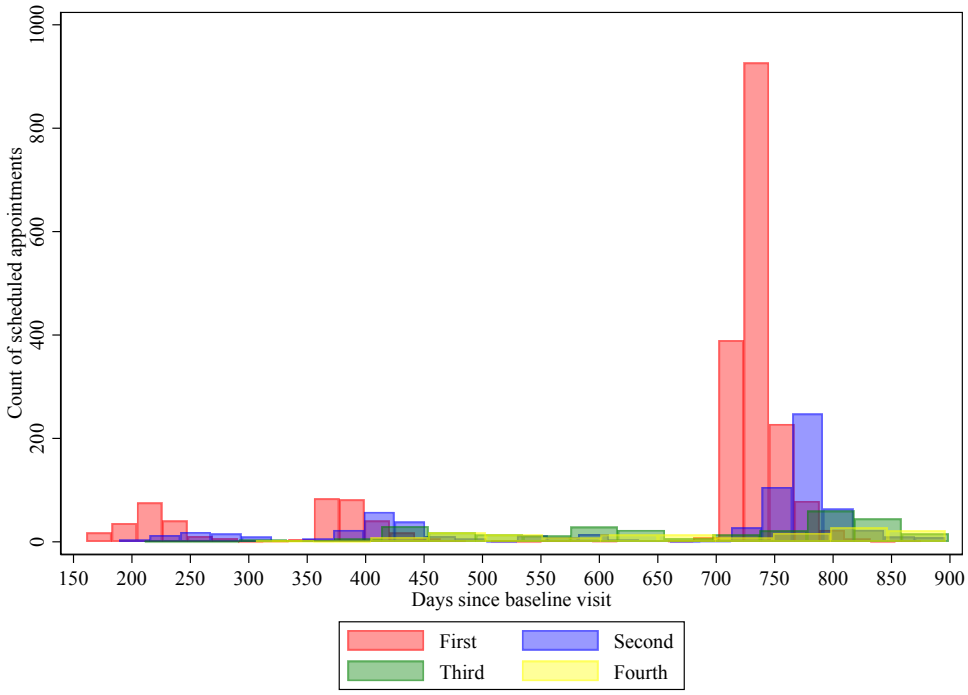
Attendance records at screening appointments were near complete, from which we assigned the results of the costing study as unit costs. Unlike the estimates of NHS reference costs, which last reported a mean unit cost of £31 in 2010/11,[10] our method allowed a more granular characterisation of costing distributions whether patients attended, did not attend, or withdrew from screening, and without any additional patient burden that would normally be associated with this kind of data collection. For withdrawals, we assigned screening costs of £0 for those 110/4536 participants who died, were pregnant, or who were no longer diabetic and as such were no longer eligible for screening. If safety standards of screening are maintained, then the costs of NHS screening represent the differential cost driver of individualised versus annual screening intervals, assuming cost effective treatment further down patient pathways. External healthcare records were additionally considered to isolate treatments attributable to DR, however only a handful of these were observed during the within trial time horizon of two years. These treatments would have introduced bias and noise into our estimates and not represent the true differences in trial arm costs. Furthermore we do not capture the costs of confirmatory slit lamp biomicroscopy following a referable screen positive event due to the procedure taking place outside of the screening pathway. As a result our analysis slightly *underestimates* the propective cost savings of individualised screening versus usual care through a reduction in screening false positives. Further work which extrapolates towards a lifetime horizon should experience no such difficulty in including these treatments.

Time horizon

Patient health state utilities were assigned to timepoints allowing for a 90 day attendance window (patients were allowed up to two additional appointment invitations if they did not attend their original scheduled appointment before reverting to usual care). However this presented difficulties at the 24 month timepoint as this window assumes that patients attended their scheduled appointments on the first day of the grace period of previous appointments. Supplementary Figure 1 and Supplementary Figure 2 showcase the distribution of the first four scheduled screening appointments for patients in the control arm and individualised arm respectively, where we observe this compounding and significant lag in scheduling, frequently for those who missed earlier appointments. The scheduling of screening appointments incurs an irreversible cost, which will be accrued whether the patient attends or not. Consequently we allowed a maximum further 90 day attendance window for patients to capture these 24 month treatment-attributable costing distributions.



Supplementary Figure 1. Control arm distributions (frequency) of scheduled screening appointments



Supplementary Figure 2. Individualised arm distributions (frequency) of scheduled screening appointments

Multiple imputation

To control for unplanned missingness, assumed to be missing at random (MAR), methods of multiple imputation were applied as they build into their model the inherent uncertainty associated with the missing data: Multiple Imputation of Chained Equations (MICE) were performed using available case data, where a separate conditional distribution was specified for each imputed variable.[11, 12] Variable selection, excluding the EQ-5D, EQ-VAS, and HUI3 which must always be included as our outcomes of interest, was resolved in part through stepwise regression on utilities (forward selection due to the number of candidate variables excluding

the possibility of backward elimination), manual testing of fit judged by Akaike Information Criterion (AIC), and logit modelling of missingness mechanisms. We then examined variable covariance, in combination with the heterogeneity of variables across the total sample, to reduce overfitting and ensure model convergence.

We observed significant collinearity of health state utilities at six and eighteen months due to smaller sample sizes within the intervention arm, and as such excluded these timepoints from our model. However, we consider it unlikely to impact our estimates of mean differences: 1. This did not affect complete case observations. 2. affected patients, at worst, matched the granularity of those observed in the control arm for the purposes of AUC QALY construction. 3. Imputing at the index rather than QALY level avoided the greater bias of excluding a larger number of available case observations.

Final model specification ran fifty imputation sets (m50),[13] and included conditional functions of gender, smoking status, age, self-completion, presence of retinopathy in either one or both eyes at baseline or first followup,[14] and the EQ-5D, EQ-VAS, and HUI3 utilities across all appropriate time points. Following methods of best practice, multiple imputation was conducted individually for control and individualised arms respectively. Predictive mean matching (k-nearest neighbours (knn) 5) was utilised for utilities due to their non-normal distribution (heavy left tails) and bounded values. As costs presented minimal missingness and converted from participant attendance history, observations for the 145 participants who moved to a non-participant GP were imputed using PMM of total screening and societal costs.[15] We observed a slight difference in incremental cost between those who informed our QoL data, and those patients across the larger distribution set of screening costs, leading us to question the validity of our MAR assumption. After examination, these differences appear to be a relic of sample size – increased noise of screening behaviours in both arms - perhaps provoked by observation at the screening clinic. We considered it prudent to impute costs across the larger set excluding QoL data (imputation of utilities did include costs as a passive predictor). Logit functions of categorical variables were augmented to avoid the bias of perfect prediction.[16] The models were run multiple times, and QALY distributions (derived from multiply imputed utilities) were visually inspected to confirm the robustness of parameter estimates.

Cost effectiveness methodology

Between arm differences in arithmetic means are the parameters of interest within economic evaluations of randomised control trials.[17] Following multiple imputation we ran seemingly unrelated regressions (SUR)[18] to estimate incremental QALYs and costs; a simultaneous method which permits baseline and covariate adjustment, while capturing the multivariate distributions of our outcomes through allowing correlation between the error terms of our regressions. To reiterate our earlier disucssion – due to possible violation of MAR we present point estimates (and CIs) of multivariate distributions of utilities, and univariate distributions of costs. All analysis followed Rubins combination rules for estimation within multiply imputed sets.[19]

These SURs were bootstrapped, a nonparametric method to characterise sampling distributions (practical when these distributions do not hold a Gaussian form), to derive bias-corrected confidence intervals around the between-arm mean differences in QALYs and costs.[17, 20] We would like to particularly recommend the paper of Schomaker and Heumann for their fantastic work on "Bootstrap inference when using multiple imputation".[21] We opted for MI BOOT (pooled sample), at the cost of greater symmetry in our confidence intervals, due to the infeasibility of BOOT MI procedures regarding the large number of imputed sets and large sample size (months rather than days of estimation time).

Mean differences in QALYs and costs were used to construct incremental cost effectiveness ratios (ICERs), and scattered on the cost effectiveness plane to represent the sampling uncertainty around means. Statistics of incremental net monetary benefit (INMB) represent the monetary value to the NHS of individualised screening intervals per patient when the willingness-to-pay threshold is known. By rearranging the decision rule, where a treatment is cost effective if the ICER is less than the threshold, a therapy should be adopted if the INMB>0. Initially we used a lower bound threshold of £20,000 of the cost effectiveness of interventions advised by the National Institute for Health and Care Excellence,[22] as if individualised screening reported a positive INMB at £20,000, any movement toward the upper threshold of £30,000 will correspond to an increase in size and not sign. CEACs present to policymakers the likelihood of making the wrong as well as right decision; where we derived individual INMB, varied across a range of thresholds, for each bootstrapped iteration of the multivariate distributions of incremental QALYs and costs. The resulting plot presents the proportion (probability) cost effective of individualised screening by willingness-to-pay thresholds for one QALY (EQ-5D/HUI3).

Within trial subgroup analyses were not conducted due to our concerns surrounding low sample sizes,[23] and that subgroupings of interest (risk category) were inherently assigned by the risk engine and so not observed in the control arm. We are developing a risk-based model in order to investigate further.

ESM RESULTS - HEALTH ECONOMICS

We report the results of the microcosting study on screening for sight-threatening diabetic retinopathy within the Liverpool Diabetic Eye Screening Programme (LDESP) in Supplementary Table 11. Within-trial summary statistics are presented in Supplementary Tables 12 and 13 for the control and individualised arms respectively.

Supplementary Table 11. Micro-costing of the Liverpool Diabetes Eye Screening Programme and estimated costs of the Risk Calculation Engine

Variable		Top-down annual total	Per attendance (n=16736)	Per non-attendance (n=6179)
Programme costs				
	Staff (including oncosts)	£242,715	£10.59	£10.59
	Stationary	£3,056	£0.13	£0·13
	IT	£22,988	£1.00	£1.00
	Total	£268,759	£11.73	£11.73
Photography				
	Staff	£157,528	£9.08	£0.91
	Cameras and equipment	£11,777	£0.68	£0.07
	Medical consumables	£4,229	£0.25	£0.00
	Total	£173,534	£10.00	£1.00
Grading (P3-P6)				
	Staff	£117,234	£7.00	£0.00
Total NHS cost		£559,528	**£28.73**	**£12.73**
Societal costs associated with screening attendance				
	Patient-borne costs	-	£2.64	-
	Productivity loss	-	£6.36	-
Total societal cost		-	**£9.00**	-

Risk calculation engine (RCE) costs		**Annual total**	**Per patient**
	Database administrator	£34,000	£1.54
	CCG administrator (20% FTE)	£6,000	£0.27
	RCE total	£40,000	**£1.81**

Not all values sum perfectly due to rounding. Estimated costs to run and maintain a data warehouse and risk engine based on Liverpool eligible population of 22,099. Excludes start-up costs.

Supplementary Table 12. Control arm summary statistics of intention-to-treat health economic data at 24 months

Variable	**Control arm** (*n*=2269)					
	n	Mean	SD/95% CI	Min.	Median	Max.
Baseline						
EQ-5D	400	0·70	0·30	-0·27	0·77	0·99
EQ-VAS	403	0·74	0·22	-0·90	0·80	1·00
HUI3	350	0·70	0·30	-0·21	0·79	1·00
6 months						
EQ-5D	9	0·00	0·00	0·00	0·00	0·00
EQ-VAS	9	0·00	0·00	0·00	0·00	0·00
HUI3	11	0·00	0·00	0·00	0·00	0·00
12 months						
EQ-5D	271	0·66	0·30	-0·29	0·73	0·99
EQ-VAS	274	0·71	0·25	0·00	0·75	1·00
HUI3	253	0·65	0·35	-0·22	0·75	1·00
18 months						
EQ-5D	57	0·44	0·41	-0·23	0·55	0·98
EQ-VAS	60	0·51	0·38	0·00	0·65	1·00
HUI3	60	0·38	0·43	-0·27	0·27	1·00
24 months						
EQ-5D	284	0·64	0·34	-0·23	0·73	0·99
EQ-VAS	288	0·69	0·27	0·00	0·75	1·00
HUI3	253	0·63	0·36	-0·27	0·75	1·00
Total QALYs						
EQ-5D						
Unadjusted	278	1·33	0·60	-0·40	1·49	1·96
Baseline adjusted	278	1·34	(1·30 to 1·37)	-	-	-
EQ-VAS						
Unadjusted	281	1·42	0·45	0·09	1·53	2·00
Baseline adjusted	281	1·40	(1·36 to 4·43)	-	-	-
HUI3						
Unadjusted	216	1·36	0·62	-0·37	1·55	2·00
Baseline adjusted	216	1·36	(1·32 to 1·41)	-	-	-
Total costs (£)						
Non-attendance	2194	9·24	(8·61 to 9·88)	0·00	0·00	76·38
Screening	2194	85·54	(84·86 to 86·25)	28·73	86·19	137·11
Societal	2194	109·44	(108·61 to 110·29)	37·73	113·19	164·11

EQ-5D = Euroqol five dimension five level instrument mapped to the three level value set, EQ-VAS = Euroqol visual analogue scale, HUI3 = Health Utilities Index Mark 3. Screening costs per patient include the costs of non-attendance. Societal costs are the combination of NHS screening costs, productivity losses, and out-of-pocket expenses. We present 95% confidence intervals in place of standard deviations for highly skewed distributions, such as costs, and for QALYs where necessitated by regression adjustment.

Supplementary Table 13: Individualised arm summary statistics of intention-to-treat health economic data at 24 months

Variable	**Individualised arm** (*n*=2265)					
	n	Mean	SD/95% CI	Min.	Median	Max.
Baseline						
EQ-5D	399	0·70	0·30	-0·20	0·77	0·99
EQ-VAS	403	0·73	0·22	0·05	0·78	1·00
HUI3	352	0·71	0·30	-0·21	0·82	1·00
6 months						
EQ-5D	35	0·38	0·43	-0·30	0·30	0·99
EQ-VAS	33	0·47	0·42	0·00	0·50	1·00
HUI3	35	0·38	0·43	-0·13	0·06	1·00
12 months						
EQ-5D	61	0·55	0·42	-0·14	0·67	0·99
EQ-VAS	60	0·60	0·35	0·00	0·70	1·00
HUI3	60	0·54	0·42	-0·15	0·72	1·00
18 months						
EQ-5D	39	0·48	0·45	-0·10	0·65	0·98
EQ-VAS	40	0·54	0·41	0·00	0·70	1·00
HUI3	39	0·47	0·44	0·00	0·55	1·00
24 months						
EQ-5D	270	0·65	0·35	-0·41	0·76	0·99
EQ-VAS	269	0·68	0·27	0·00	0·75	1·00
HUI3	226	0·63	0·37	-0·28	0·78	1·00
Total QALYs						
EQ-5D						
Unadjusted	261	1·35	0·62	-0·60	1·55	1·97
Baseline adjusted	261	1·34	(1·30 to 1·38)	-	-	-
EQ-VAS						
Unadjusted	267	1·38	0·48	0·05	1·55	2·00
Baseline adjusted	267	1·40	1·37 to 1·44	-	-	-
HUI3						
Unadjusted	192	1·35	0·66	-0·34	1·64	2·00
Baseline adjusted	192	1·35	(1·30 to 1·39)	-	-	-
Total costs (£)						
Non-attendance	2195	6·77	(6·20 to 7·39)	0·00	0·00	101·84
Screening	2195	68·10	(67·15 to 69·05)	32·35	61·08	207·65
Societal	2195	86·18	(84·99 to 87·31)	41·35	79·08	243·65

EQ-5D = Euroqol five dimension five level instrument mapped to the three level value set, EQ-VAS = Euroqol visual analogue scale, HUI3 = Health Utilities Index Mark 3. Screening costs per patient include the costs of non-attendance. Societal costs are the combination of NHS screening costs, productivity losses, and out-of-pocket expenses. We present 95% confidence intervals in place of standard deviations for highly skewed distributions, such as costs, and for QALYs where necessitated by regression adjustment.

At a two year time horizon we report a small number of negative QALYs. This is a drawback of health state utility tariffs utilising a scale of dead to perfect health,[24] where a select few participants reported health state utilities worse than death (<0) at each time point, producing negative QALYs. This seemed feasible given the elderly demographics of trial participants and the likelihood of multi-morbidity, and was uniformly balanced between trial arms.

As expected, all ICER point estimates fell in the south east and south west quadrants signifying the dominance of individualised screening in cost savings. Mean differences in QALYs clustered around the zero threshold across all instruments. Estimates of joint distributions following multiple imputation supported no difference or dominance in QALYs and cost savings (screening, societal). While reducing screening frequency may well improve patient quality of life, we doubt our applied instruments are sensitive enough to pick up this effect. Similarly we consider it unlikely that our time horizon is long enough to observe the benefits to those high risk patients who were confirmed STDR positive and as such received treatment sooner than they would have under annual screening. The dominance in QoL judged by the EQ-5D and HUI3 may be a remnant of the poor suitability of models of multiple imputation in predicting small marginal effects when missingness is high.[13]

Discussion

Unlike traditional cost effectiveness analysis, individualised screening was not strictly expected to provide an increase/decrease in QALYs for an increase/decrease in cost, where the stated ICER is typically interpreted as the cost to the NHS of producing one QALY. Considering the slow progression of the development of diabetic retinopathy, and that frequency of screening appointments would plausibly have little effect on participants' health state utilities when holding safety standards constant, the observation of equivalence (or rather no skew towards QALY loss/patient across a range of sensitivity adjustments) is both expected and encouraging in viewing individualised screening as a cost minimiser. This is supported by the agreement between our multiply imputed and observed complete case utility data, and secondary safety outcomes of the ISDR trial.

This study well characterised the costs of NHS screening and treatment adherence of both standard and individualised practice that would affect screening populations such as in Liverpool. While intended to reduce patient burden, it would have been useful to collect utility data for the entire cohort. This within trial economic evaluation operated at a two year time horizon, and the effects of screening intervals on quality of life appear to be close to negligible, where our between-arm utilities and incremental QALYs demonstrated near-equivalence. A longer time horizon would be necessary to capture the possible effects on participant QoL and include blind years averted. Similarly we may not capture lifetime costing differentials (our assumptions tended to the conservative and as such may underestimate cost savings), therefore future research which fully capture diabetic retinopathy attributable treatments is advisable. Our findings provide evidence towards the characterisation of utility and costing distributions for the synthesis of evidence in health technology assessment (HTA). A risk-based model is in development to extrapolate across a lifetime horizon.

Supplementary references

1. Machin D, Campbell M, Fayers P, Pinol A. Sample Size Tables for Clinical Studies. Blackwell Science, 3rd edition, Oxford. 1997
2. Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2002; 2nd edition, New York: John Wiley.
3. Steyerberg EW, van Veen M. Letter: Imputation is beneficial for handling missing data in predictive models. J Clin Epidemiol. 2007;60:979.
4. National Institute for Health and Care Excellence. Position statement on use of the EQ-5D-5L value set for England (updated October 2019). 2019; https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l
5. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the economic evaluation of health care programmes. 2015; 4th ed. Oxford University Press, Oxford.
6. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Economics, 2005; 14(5), 487-496.
7. Gabrio A, Hunter R, Mason A, Baio G. Pitfalls of adjusting for mean baseline utilities/costs in trial-based cost-effectiveness analysis with missing data. 2018; arXiv:1805.07149.
8. Curtis L, Burns A. Unit Costs of Health and Social Care 2017. Personal Social Services Research Unit, University of Kent, Canterbury; 2017. https://doi.org/10.22024/UniKent/01.02/65559.
9. Office for National Statistics. Dataset: Consumer price inflation time series. 2019; https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceindices
10. Department of Health and Social Care. NSRC4 NHS Trusts and PCTs Combined Reference Cost Schedules 2010-11. 2011; https://www.gov.uk/government/ publications/2010-11-reference-costs-publication
11. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Statistics in Medicine 2011; 30, 37.
12. Royston P, White IR. Multiple imputation by chained equations (MICE): implementation in Stata. J Stat Softw 2011; 45(4), 1-20.
13. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007; 8: 206-213.
14. Stratton IM, Aldington SJ, Taylor DJ, Adler AI, Scanlon PH. A simple risk stratification for time to development of sight-threatening diabetic retinopathy. Diabetes Care. 2013 Mar;36(3):580-5. doi: 10.2337/dc12-0625.
15. MacNeil Vroomen J, Eekhout I, Dijkgraaf MG, et al. Multiple imputation strategies for zero-inflated cost data in economic evaluations: which method works best?. Eur J Health Econ 17, 939–950 (2016). https://doi.org/10.1007/s10198-015-0734-5
16. White IR., Daniel RM, Royston P. Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical data. 2010; Computational Statistics & Data Analysis 54: 2267–2275.
17. Glick A, Doshi JA, Sonnad SS, Polsky D. Economic evaluation in clinical trials. 2015; 2nd ed. Oxford university press, Oxford.
18. Davidson, R; MacKinnon, JG. Estimation and inference in econometrics. 1993; Oxford University Press. ISBN 978-0-19-506011-9.
19. Rubin, DB. Multiple imputation after 18+ years. 1996; Journal of the American Statistical Association 91: 473–489.
20. Briggs A, Wonderling D, Mooney, C. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Economics 1997; 6:327–340.
21. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation.Statistics in Medicine. 2018; 37:2252–2266.https://doi.org/10.1002/sim.7654
22. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013; NICE article [PMG9]. https://www.nice.org.uk/article/pmg9/chapter/Foreword
23. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. Health Technol Assess 2001;5(33).
24. Furlong W, Feeny D, Torrance G, et al. Multiplicative Multi-Attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report. Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, Canada; 1998.

ISDR Trial

Introducing personalised risk based intervals in screening for diabetic retinopathy: development, implementation and assessment of safety, cost-effectiveness and patient experience. Workstream E: Randomised trial comparing standard and test screening intervals protocols

Version 6.0

20th November 2018

Study Sponsors:
University of Liverpool
Research Support Office
2nd Floor Block D Waterhouse Building
3 Brownlow Street
Liverpool
L69 3GA

Research and Innovation Department
Royal Liverpool and Broadgreen University Hospitals Trust
Prescot Street
Liverpool
L7 8XP

Sponsor references:
Trust RD&I: 4660
University of Liverpool: UoL000994

ISRCTN number: 87561257

Other:

Protocol Approval

Authorised by Chief Investigator:

Signature: _____ **Date:** _____

Prof Simon Harding
Department of Eye and Vision Science
University of Liverpool
3rd floor UCD Building
Daulby Street
Liverpool
L69 5QA

Tel: 0151 706 4070/3966
E-mail: s.p.harding@liverpool.ac.uk

Authorised on behalf of the Sponsors:

Signature: _____ **Date:** _____

Mr Alex Astor
Research Support Manager
2nd Floor Block D Waterhouse Building
3 Brownlow Street
University of Liverpool
Liverpool L69 3GL
Tel: 0151 794 8739
Email: Alexander.Astor@liverpool.ac.uk
sponsor@liverpool.ac.uk

Authorised by Lead Statistician:

Signature: _____ **Date:** _____

Dr. Marta García-Fiñana
Reader in Biostatistics
Department of Biostatistics
Block F, Waterhouse Bld
1-5 Brownlow Street
Liverpool,
L69 3GL

Tel: 0151 794 9755
E-mail: martaf@liv.ac.uk

General Information

This randomised controlled trial is part of a five year Programme of Applied Research awarded by NIHR, ref RP-PG-1210-12016 (1st Feb 2013 – 31st Jan 2018). It follows on from a development project that was funded by NIHR as a Programme Development Grant (NIHR ref RP-DG-0709-10138, Trust RD&I No. 4102, UOL: 000712, Oct 2010 – 30th Sep 2011).

The full programme of research has been split across three separate sponsorship and REC applications; each application contains components of the programme matched for ethical and governance issues. This protocol describes Workstream E - the third component of this Programme of Applied Research. It will commence in year 2 of the 5 year programme. Further information is available in Background Information, Section 2 of this protocol.

The protocol should not be used as an aide-memoir or guide for the management of other patients; every care was taken in its drafting, but corrections or amendments may be necessary. These will be circulated to the registered investigators in the trial. Clinical problems relating to this trial should be referred to the relevant Chief Investigator via the Clinical Trials Research Centre (CTRC).

This protocol defines the participant characteristics required for study entry and the schedule of treatment and follow-up. Participant recruitment will be undertaken in compliance with this document and applicable regulatory and governance requirements and waivers to authorise non-compliance are not permitted.

Incidence of protocol non-compliance, whether reported prospectively or retrospectively noted (e.g. as a result of central monitoring) are recorded as protocol deviations, the incidences of which are monitored and reported to trial oversight committees.

Statement of Compliance

This study will be carried out in accordance with the World Medical Association Declaration of Helsinki (1964) and the Tokyo (1975), Venice (1983), Hong Kong (1989) and South Africa (1996) amendments and will be conducted in compliance with the protocol, CTU Standard Operating Procedures.

This trial does not involve the use of a medicinal product or medical device. Therefore this trial is not considered to be a clinical investigation under EU Directive 2001/20/EC. The intervention is a change to the screen interval based on an estimate of individual risk. The trial will adhere to standards for Good Clinical Practice in accordance with ICH-GCP guidance and to the Research Governance Framework.

Relationship Statements

The UK Clinical Research Collaboration (UKCRC; www.ukcrc.org) is a partnership organisation working to establish the UK as a world leader in clinical research. Following a review by an international panel, the Clinical Trials Research Centre (CTRC) at the University of Liverpool has been assessed as reaching the highest quality standard required by the UKCRC and achieved full UKCRC registration.

The CTRC encompasses clinical trials activity in areas including medicines for children (The Medicines for Children Research Network Clinical Trials Unit; MCRN CTU), cancer (The Liverpool Cancer Trials Unit; LCTU), epilepsy, oral health and obstetrics and gynaecology (http://www.ctrc.org.uk/). All CTRC activities are underpinned by methodological rigour, a

modern data management system, similar technical requirements and a common set of standard operating procedures.

The NIHR Medicines for Children Research Network and National Cancer Research Network is part of the National Institute for Health Research Clinical Research Network.

Contact Details: Institutions

Sponsor:	Trial Management and Monitoring:	Clinical Laboratory:
University of Liverpool Research Support Office 1st floor Duncan Building Daulby Street Liverpool L69 3GA Tel: 0151 794 8373 Research Development and Innovation Department Royal Liverpool and Broadgreen University Hospitals Trust Prescot Street Liverpool L7 8XP Tel: 0151 706 3702	Clinical Trials Research Centre University of Liverpool MC Clinical Trials Unit Institute of Child Health Alder Hey Children's NHS Foundation Trust Liverpool L12 2AP Tel: 0151 282 4724 Fax: 0151 282 4721	N/A
Other Medical or Technical Departments and/or Institutions (1):	**Other Medical or Technical Departments and/or Institutions (2):**	**Other Medical or Technical Departments and/or Institutions (3):**
Liverpool Diabetic Eye Screening Programme 3rd Floor UCD Building Royal Liverpool and Broadgreen University Hospitals Trust Prescot Street Liverpool L7 8XP 0151 706 5784/5	Department of Eye and Vision Science Institute of Ageing and Chronic Disease Faculty of Health and Life Sciences University of Liverpool William Henry Duncan (Apex) Building 6 West Derby St Liverpool L7 8TX	

Contact Details: Individuals

Individual Authorised to Sign the Protocol and Protocol Amendments on behalf of the Sponsor:	**Chief Investigator (CI):**
Mr Alex Astor Research Support Manager 2nd Floor Block D Waterhouse Building 3 Brownlow Street University of Liverpool Liverpool L69 3GL Tel: 0151 794 8739 Email: Alexander.Astor@liverpool.ac.uk sponsor@liverpool.ac.uk	Prof Simon P Harding Professor of Clinical Ophthalmology Head, Department of Eye and Vision Science Honorary Consultant Ophthalmic Surgeon Department of Eye and Vision Science Institute of Ageing and Chronic Disease Faculty of Health and Life Sciences University of Liverpool William Henry Duncan (Apex) Building 6 West Derby St Liverpool L7 8TX University enquiries 0151 794 9051 NHS Secretary Helen Flinn 0151 706 3966 Email: s.p.harding@liverpool.ac.uk
Medical Expert who will Advise on Protocol Related Clinical Queries (If other than CI):	
Dr Deborah M Broadbent, MRCOphth Honorary Senior Lecturer, Department of Eye and Vision Science 3rd Floor UCD RLBUHT Prescot St Liverpool L7 8XP Tel: 0151 706 2408 Fax: 0151 706 2419 Email: deborah.broadbent@rlbuht.nhs.uk	

The contact details of other individuals involved in the trial are detailed in the Trial Oversight Committee Membership document located in the Trial Master File.

Table of Contents

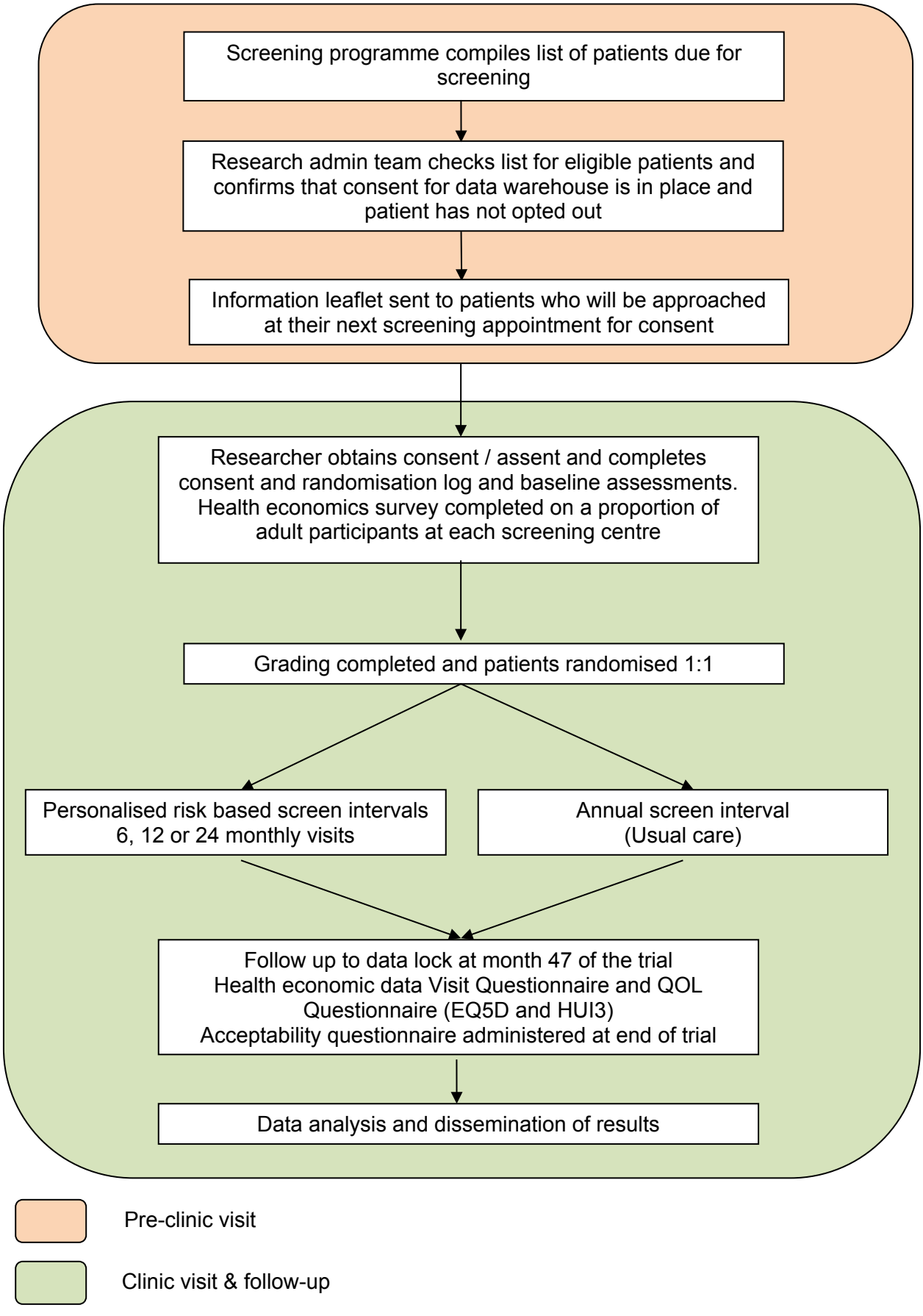
Glossary

AE	Adverse Event
BDR	Background diabetic retinopathy
CCG	Clinical Commissioning Group
CI	Chief Investigator
CRF	Case Report Form
CTRC	Clinical Trials Research Centre
DM	Diabetic maculopathy
DR	Diabetic retinopathy
GP	General Practitioner
GUI	Graphical User Interface
HES	Hospital Eye Service
IB	Investigator's Brochure
IDSMC	Independent Data and Safety Monitoring Committee
IEC	Independent Ethical Committee
LDESP	Liverpool (National) Diabetic Eye Screening Programme
LMC	Local Medical Committee
NDESP	NHS Diabetic Eye Screening Programme
NIHR	National Institute for Health Research
NIHR CRN	National Institute for Health Research Clinical Research Network
PI	Principal Investigator
PPDR	Preproliferative diabetic retinopathy
PPI	Patient and Public Involvement
PWD	Person with diabetes
R&D	Research & Development
RCE	Risk Calculation Engine
RCT	Randomised controlled trial
REC	Research Ethics Committee
RN	Research Nurse
SAE	Serious Adverse Event
SDV	Source Data Verification
SPEU	St. Paul's Eye Unit
STDR	Sight-threatening Diabetic Retinopathy
TMG	Trial Management Group
TSC	Trial Steering Committee
VA	Visual acuity
VI	Visual impairment

1 PROTOCOL SUMMARY

Title:	ISDR: Individualised Screening for Diabetic Retinopathy Introducing personalised risk based intervals in screening for diabetic retinopathy: development, implementation and assessment of safety, cost-effectiveness and patient experience. Workstream E - Randomised trial comparing standard and test screening intervals protocols
Phase:	Not applicable
Population:	This trial aims to recruit around 4460 (see 9.2) people with diabetes under the care of Liverpool CCG who are undergoing screening for diabetic retinopathy.
Study Centres and Distribution:	Single site - Royal Liverpool University Hospital and 6 community centres across Liverpool providing screening for diabetic retinopathy in the Liverpool Diabetic Eye Screening Programme (LDESP)
Study Duration:	54 months (19 months recruitment, 24 months follow up, 3 months to allow for a 90 day window during follow-up and 8 months for data cleaning and statistical analysis)
Primary Programme Objective:	To develop and implement a major enhancement to the existing delivery of screening for sight threatening diabetic retinopathy (STDR) by introducing an individualised approach based on measured patient-centred risk, acceptability to patients and staff, safety and cost-effectiveness.
Primary Trial Objective:	To evaluate the safety and acceptability of variable interval screening in a pragmatic whole population setting as measured by attendance rates to screening.
Secondary Trial Objectives:	To assess for personalised risk based interval compared to annual screening: rates and severity of DR; visual outcomes; impact on general diabetes care; Cost effectiveness and quality of life.

Schematic of trial design:



2 BACKGROUND INFORMATION

2.1 Introduction

A 5 year National Institute for Health Research (NIHR) Programme of Applied Research was awarded to the Royal Liverpool University Hospitals NHS Trust in February 2013: "Introducing personalised risk based intervals in screening for diabetic retinopathy: development, implementation and assessment of safety, cost-effectiveness and patient experience." Harding SP, Broadbent DM, Gabbay M, Grey P, James M, Stratton I, Fisher AC, Vora JP, Roberts J, Byrne P, Garcia-Finana M, Breen R, Williamson P. (RP-DG-1210-12016).

This Programme of Applied Research comprises the following work streams:
Workstream A: Systematic review
Workstream B: Prospective observational study of risk factors for progression to sight threatening diabetic retinopathy (STDR), treatment and visual impairment (VI)*
Workstream C: Risk calculation engine development & testing*
Workstream D: Health economics
Workstream E: Randomised trial comparing standard and test screening intervals protocol
Workstream F: Exploring perceptions of screening and variable screening intervals amongst people with diabetes (PWD) and professionals
Workstream G: Knowledge transfer and preparation for implementation

It follows on from a 12 month NIHR Programme Development project that was conducted in 2011 (Trust RD&I ref 4202):
"Acceptability and effectiveness of risk based intervals in screening for diabetic retinopathy – towards a personalised approach. Harding SP, Gabbay M, Grey P, James M, Stratton I, Broadbent M, Fisher A, Vora J, Roberts J, Byrne P, Garcia-Finana M. (RP-DG-0709-10138)

This protocol describes the randomised controlled trial in Workstream E.

***Ethical approval for Workstreams B (data warehouse) and C (Risk calculation engine) has already been granted by North Preston REC – reference 13/NW/0196**

2.2 Diabetic retinopathy

In 2009 around 2.6 million people in the UK had been diagnosed with diabetes mellitus (DM), and it is estimated that by 2025 there will be more than four million [1]. DM is a life-long condition associated with the development of various health complications, including diabetic retinopathy (DR), a progressive condition affecting the retina, which can lead to visual impairment and blindness. After 20 years with diabetes, nearly all patients with Type 1 diabetes [2] and over 60% with Type 2 diabetes [3] will have some degree of DR. DR affects people of all ages and used to be the leading cause of blindness in people of working age in the UK. However Liew et al reviewed the causes of blindness certifications in England and Wales in 2009-2010 and found that DR had dropped to the second commonest cause compared to ten years earlier [4]. Although many factors will have contributed to the results it is likely that screening for DR has played a key role. The number of people worldwide with diabetes quadrupled between 1980 and 2014 [5].

The early stages of DR (background and mild pre-proliferative retinopathy) are asymptomatic and do not require any treatment. Sight threatening diabetic retinopathy or STDR (an umbrella term that encompasses sight threatening levels of retinopathy and maculopathy) may still be asymptomatic but requires close monitoring by a hospital eye service (HES), and sometimes treatment, to prevent visual impairment (VI). Patients with screen positive DR are referred to the Hospital Eye Service for assessment by an ophthalmologist.

Treatment (by laser photocoagulation or injections into the eye) aims to stop progression and stabilise the retinal changes in order to preserve sight. Treatment cannot always reverse the process once vision is lost. Screening is therefore recommended to identify STDR before any loss of vision is noticed.

2.3 Screening for diabetic retinopathy

In the NHS Diabetic Eye Screening Programme (NDESP) 68 local programmes undertake primary care-based screening. At screening appointments patients have their visual acuity (VA) checked using logMAR charts and then high quality digital photographs are taken of the back of their eyes (their retinas) [6]. The photographic images are later graded in the grading centre by accredited technicians. Patients who screen positive for STDR are referred to their local HES for further evaluation.

Systematic screening for DR was introduced across Britain by 2007 and annual screening for all people with diabetes aged 12 years or older was recommended. The lower age limit is due to diabetic retinopathy being rare before puberty. There is no upper age limit. The screen interval was based on expert opinion, rather than direct evidence.

Evidence shows that many people are at low risk of developing referable (screen positive) DR between annual screening appointments and could safely be screened less often [7], whilst others are at high risk and might benefit from more frequent screening [8].
A large cohort study (n=4770, 20,570 screen visits) conducted in Liverpool in the 1990s, The Liverpool Diabetic Eye Study (LDES) showed that people with Type 2 DM and no existing DR have a very low risk (0.3%) of developing STDR within a one year period [9], while people who already had some retinopathy had a higher risk. The table below shows the cumulative incidences of STDR for people with different levels of DR at their baseline screening appointment in patients with Type 2 DM.

Baseline retinopathy level	**Cumulative incidence of STDR at:**		
	1 year	*2 years*	*5 years*
No retinopathy (LDES level 10)	0.3%	0.8%	3.9%
Background retinopathy (LDES level 20)	5.0%	11.2%	28.9%
Mild pre-proliferative retinopathy (LDES level 30)	15.0%	27.8%	63.2%

Similar results were found in patients with Type 1 DM [10].

According to this data, the majority of people attending screening programmes are at low risk based on their current retinopathy levels.

In the same cohort of patients from Liverpool [11] the overall prevalence of any retinopathy, STDR and proliferative diabetic retinopathy (PDR) in patients with Type 1 DM was 45.7%, 16.4% and 3.7%, respectively. No retinopathy was observed in 53.2%, background diabetic

retinopathy (BDR) in 26.2%, mild pre-proliferative DR (PPDR) in 9.9%, moderate PPDR (referable) in 4.0% and sight-threatening maculopathy (referable) in 12.3%.

The prevalence of any retinopathy, STDR and PDR in patients with Type 2 DM was 25.3%, 6.0% and 0.5%, respectively. The prevalence of BDR, mild PPDR and moderate PPDR was 17.7%, 5.2% and 1.6%, respectively. Sight-threatening maculopathy was present in 5.1% of patients.

Another UK cohort study (UKPDS) followed newly diagnosed patients with Type 2 DM for up to nine years. Of the 2316 patients who had no retinopathy at baseline, only 0.3% required laser photocoagulation at three years, 1.1% at six years and 2.6% at nine years [12]. Of those with pre-proliferative or proliferative retinopathy at baseline, 6.6% needed treatment after three years and 13.3% after nine years.

2.3.1 **Updates in Screening Practice**

In 2016 the UK National Screening Committee (NSC) changed its recommendation to say that people with diabetes at low risk of sight loss could have their screen interval increased from one to two years but this has not been widely adopted [13]. They suggested a stratified screening approach where patients with no diabetic retinopathy (DR) on two successive screen visits would be classified as at low risk of developing sight threatening diabetic retinopathy (STDR). This was based on the evidence from a large observational study and cost-effectiveness analysis [14].

2.4 Risk factors

In other studies, the risk of development and progression of DR to a level requiring treatment has been shown to be related to age, gender, duration of diabetes [2, 3], severity of retinopathy [12], blood HbA1c levels [15-20], blood pressure 21-23] blood lipid levels [24] and proteinuria [25].

If the relative contribution of each risk factor to overall risk could be calculated, and data on each risk factor data were available for individuals, individual risks could be estimated and screening frequencies could be tailored to the level of risk.

In Iceland a ‘risk engine’ has been produced to estimate the risk of development of STDR, based on epidemiological data on risk factors for DR (www.risk.is) [26]. Using eight demographic, systemic and ocular risk factors the risk calculator predicts the risk of developing STDR at 1 year, 5 years and 10 years. This was tested empirically on a database of Danish patients. Screen intervals between 6 and 60 months were calculated for individual patients. The mean recommended screening interval was 29 months. The number of screening visits was reduced by 59% when compared to annual screening. Using a ROC curve they found that the probability that a randomly selected patient who developed STDR would be given a higher risk score than a randomly selected patient who did not develop STDR was 76%.The algorithm predicted risk of STDR reasonably well, although it over-estimated risk in high risk groups. There are no risk engines for DR that are derived using data from a real and recent cohort of patients.

In Liverpool a generalised linear model classifier (GLM) was developed as part of our Programme Development Grant and used to test non-ophthalmic known risk factors in 541 PWD attending a hospital based assessment clinic. A graphical web-based implementation

was developed and is viewable at www.liverpooleye.org section DR_Risk_Expert_System. This was refined to include longitudinal and cross-sectional data from 11806 patients in 77/95 GP surgeries in Liverpool. The final risk model used is a Markov model and includes imputation for missing data. Collection of data from the GP surgeries forms part of Workstreams B, C and D of this programme grant. The risk stratification algorithms and the output (the ISDR Risk Calculation Engine, RCE) will be used to inform the variable interval policy to be tested in the RCT. The acceptability of the risk of missing STDR has been reviewed by the patient representatives and co-applicants and has been set at up to 2.5%. The output of the risk engine will be the calculation for each individual of the risk of developing screen positive diabetic retinopathy by given time points (6, 12 and 24 months). The Liverpool RCE used will have the capability to be continually refined and updated as new data is added but it will remain fixed for the duration of the trial, so that a consistent algorithm is used to assign screening intervals for all trial participants.

2.5 Personalised risk based screen intervals

Previous studies have shown that screening for DR is a highly cost-effective intervention [27,28]. The higher the take-up rates for screening, the higher the cost-effectiveness [29]. Logically, screening low-risk patients at two-yearly intervals should be more cost-effective than annual intervals, although it may be discovered that this cost is offset by the need to screen high risk patients more frequently than annually.

A simulation study undertaken in 1996 used the results of published epidemiological studies [[30] to compare screening policies. They estimated that where screening was accurate there was very little difference in outcome in terms of years of sight saved between a policy of annual screening until retinopathy was detected and six monthly thereafter and a policy of screening once every two years until retinopathy is detected and annually thereafter.

There is some evidence indicating that extending the screen interval in low risk patients may be possible. Data on 10-year incidence from the LDES in a population of patients with type 2 diabetes enrolled in a systematic screening programme suggested that a 3-yearly screening interval could safely be adopted for patients with no retinopathy at baseline, but yearly or more frequent screening was needed for patients with higher grades of retinopathy or insulin use [9] . Similar results were shown for Type 1 DM [10]. In Sweden biennial screen intervals have been adopted for subjects without retinopathy for some time [31]. A study carried out in Malmo prospectively followed patients with Type 2 DM and no retinopathy and concluded that it appeared safe to adopt 3-year intervals as suggested by the Liverpool group [32] However this group of patients were compliant (only 9% did not attend for follow up), had a short duration of diabetes (6 ± 6 years) and good control (HbA1c 6.4 ± 1.5% at baseline and 6.3 ± 1.3% at 3 year follow up) and consequently it would be unwise to recommend 3-year intervals for all patients with diabetes without further studies.

Two recent studies have shown that the risk of progression to STDR is significantly higher for those with BDR in both eyes than those with BDR in only one or in neither eye [33, 34]. The first study also suggested that combining the results from 2 consecutive years of photographic screening enables estimation of the risk of future development of STDR [33].

An economic model based on the UKPDS study [35] showed that screening and treatment was effective in reducing blindness and that the more infrequent the screening (from annually to once every five years) the more cost effective it was. However for low risk patients annual screening conferred little benefit above that achieved with screening every

second or third year. In addition, the cost per QALY gained was significantly higher for screening every year compared with screening less often, especially for the low risk groups.

There has been very little research into patients' understanding of screening and their views about introducing variable screening intervals. Yeo et al (2012) found that extended intervals may be acceptable to the majority of people with diabetes if there was adequate evidence to support such a change [36]. Formal patient representative feedback on the trial proposal suggested that patients might be resistant to extending screening intervals beyond 12 months, because of the reassurance that annual screening brought. However, preliminary data from semi-structured interviews with patients in Liverpool suggests that the majority were not unduly concerned with some welcoming the changes, provided that sufficient information was provided to them to explain the new system.

2.6 Summary

There is evidence that screening for DR at intervals based on individual risk would be both safe and cost-effective compared with annual screening, and from our Programme Development Grant work we have preliminary evidence that variable intervals will be acceptable for the majority of patients if supported with good quality patient information.

A data warehouse (reference 13/NW/0196) has been established to collate and link individual risk factor data from GP practices, STDR outcome data from the DR assessment clinic at St Pauls Eye Unit and the Liverpool DESP (LDESP) screening software (OptoMize). This data source has been used to develop a locally applicable 'risk engine' (reference 13/NW/0196). A number of risk factors for developing sight threatening diabetic retinopathy have been identified in the published literature and in preliminary work in Liverpool. These risk factors are both modifiable risks (e.g. blood pressure and blood sugar); and non-modifiable risks (e.g. duration since diagnosis and age). The ISDR 'risk engine' utilises a restricted number of risk factors, known as covariates, in the estimation of the risk of developing STDR. The potential covariates for which there was evidence of being informative in the risk engine prepared by the ISDR investigators were reviewed by the ISDR PPI group and potential additional covariates suggested. The covariates considered were as follows: age; duration of diabetes; insulin use; systolic blood pressure; diastolic blood pressure; total cholesterol; LDL cholesterol; current retinopathy level in both eyes; previous retinopathy level in both eyes; eGFR; HbA1c; attendance / concordance; gender; ethnicity; smoking; hospitalisations; drugs / medications; other illness / injuries; diabetic complications; major life event; involvement with other agencies; family history of diabetes; family history of visual impairment; number of other heath related issues since last screen; time abroad. The most predictive factors were chosen and these were disease state (current retinopathy levels in both eyes), age, glycated haemoglobin (HbA1c), duration of diabetes, systolic blood pressure and total cholesterol.

Our PPI group defined the degree of risk (up to 2.5% risk of developing screen positive diabetic retinopathy) acceptable to people with diabetes, allowing assignment of individuals to screen intervals at the time points 6, 12 and 24 months. The choice of these intervals has been based on a review of the available literature, our data and clinician/PPI consensus. Following each negative screening outcome, patients are assigned to the longest recall period at which their risk estimation would not exceed the 2.5% threshold.

A randomised controlled trial is needed to assess the safety, cost-effectiveness and acceptability of variable screening intervals in practice. The current trial will evaluate risk-

based screening recall intervals assigned using the Liverpool Risk Calculation Engine compared with usual fixed annual recall intervals within the Liverpool screening population.

2.7 Objectives

To evaluate the safety and acceptability of variable risk based interval screening in an unselected whole population setting.

2.8 Potential risks and benefits

2.8.1 **Potential risks**

There are no known risks to personalised risk based screening intervals.

The potential risks to participants, which will be assessed within the trial, are listed below:

- Implementing personalised risk based screening intervals may adversely affect attendance to screening for DR.
- STDR may be detected later than would have been the case with fixed annual screening intervals in people who are apparently at low risk and have an extended screening interval. This might lead to the need for more numerous interventions (laser treatment or intravitreal injections) or referral for intraocular surgery (vitrectomy) and poorer outcomes in terms of visual impairment.
- Implementing personalised risk based screening intervals may cause anxiety in patients who are accustomed to annual screening and who have their screening interval reduced or extended.
- Patients whose screening intervals are extended may not put as much effort into diabetes control and consequently be at increased risk of heart attack, stroke and kidney failure.

There is currently no evidence to support these potential risks. The potential risks were explored with our PPI group who expressed the above concerns.

2.8.2 **Known potential benefits**

The benefits of personalised risk based screening intervals have not been previously tested. The expected benefits, which will be assessed within the trial, include:

- Reduced burden of appointments for people who are at low risk of developing STDR.
- Reduced anxiety for people with diabetes with reassurance that personalised risk is being determined.
- Earlier detection of STDR in people at high risk associated with better outcomes in terms of vision and need for fewer and less invasive treatments.
- Improved cost-effectiveness through redirection of resource of the screening programme, including increased capacity to see newly diagnosed patients.

3 SELECTION OF CENTRES/CLINICIANS

The study centre is the Royal Liverpool University Hospital.

The study will be based in the Liverpool Diabetic Eye Screening Programme (LDESP) on the 3rd Floor of the University Clinical Departments in the Royal Liverpool University Hospital.

Screening takes place in 7 fixed sites across Liverpool:

- Breeze Hill Neighbourhood Medical Centre
- Everton Road Health Centre
- Fiveways Family Health Centre
- Picton Neighbourhood Health and Children's Centre
- South Liverpool Treatment Centre
- Yew Tree Centre
- Clinical Eye Research Centre, Royal Liverpool University Hospital

The treatment centre is St Paul's Eye Unit at the Royal Liverpool University Hospital.

The study centre will be initiated once all global (e.g. local R&D approval) and study-specific conditions (e.g. training requirements) have been met, and all necessary documents have been returned to CTRC. Initiation meetings will cover the requirements outlined in CTRC SOPs TM017 and TM018.

Adherence to the protocol procedures will be monitored throughout the trial by the Trial Coordinator, Data Manager and Trial Statistician. Participating centres will be expected to each maintain a file of essential trial documentation (Investigator Site File), which will be provided by the CTRC, and keep copies of all completed Case Report Forms (CRFs) for the trial. Data collection will use a combination of paper CRFs (with no carbon copies) and electronic data (see section 13.3 for details on the data capture methods).

3.1 Centre/Clinician inclusion criteria

a) Local R&D approval
b) Completion and return of 'Signature and Delegation Log' to CTRC
c) Signed contract between site and sponsor
d) Receipt of evidence of completion of (a), (b) and (c) by CTRC

3.2 Centre/Clinician exclusion criteria

Not meeting the inclusion criteria listed above

4 TRIAL DESIGN

4.1 Primary endpoint

The primary outcome will be a comparison of attendance rates for follow-up screening in the two arms of the study.

Non-attendance is defined as failure to attend any screening appointments within 90 days of the expected follow-up date, irrespective of how many appointments they have had.

4.2 Secondary endpoints

- Number of cases of STDR detected
- Retinopathy level at screening (LDES and NDESP grading)
- Maculopathy level at screening (LDES and NDESP grading)
- Number of false positive screening episodes
- Number of screening appointments
- Number of dedicated diabetes assessment clinic appointments (biomicroscopy)
- Number of other eye appointments for DR
- Visual acuity (logMAR)
- New visual impairment (using the following two thresholds: ≥ +0.50 logMAR and ≥ +0.30 logMAR in the best eye.)
- New visual impairment due to diabetic retinopathy (using the following two thresholds: ≥ +0.50 logMAR and ≥ +0.30 logMAR in the best eye)
- Number of missed appointments to screening
- Quality-adjusted life years (QALYs) estimated using EQ-5D-5L and Health Utilities Index Mark 3 (HUI3)*
- Cost per QALY gained*

*Outcomes related to health economics.

5 STUDY POPULATION

5.1 Inclusion criteria

Patients who:

I. are aged 12 years and over
II. attend for retinal screening in community clinics or screening assessment clinic during the recruitment period
III. are registered with a participating GP practice
IV. are included in the study data warehouse (have not opted out)
V. have no retinopathy or have retinopathy and maculopathy less than the definition of screen positive DR
VI. have gradable digital retinal images in both eyes
VII. give their informed consent for participation
VIII. are not involved in any trial investigating a treatment aiming at preventing or modifying the development of STDR

5.2 Exclusion criteria

Patients who:

I. are under 12 years
II. are not registered with a participating GP practice
III. have opted out from the study data warehouse
IV. have screen positive DR
V. have significant other eye disease requiring referral to the HES
VI. are ineligible for screening for whatever reason, including having ungradable digital retinal images, including patients who have only one eye or an ungradable eye with no visual potential
VII. do not give consent for participation in the RCT
VIII. are involved in any trial investigating a treatment aiming at preventing or modifying the development of STDR

5.3 Participant transfer, premature discontinuation of trial intervention and withdrawal

In consenting to the trial participants are consenting to randomised trial intervention, follow-up and data collection. If a participant requests to be withdrawn from the trial at any point they will revert to routine care. Reasons for withdrawal will be collected on a withdrawal / loss to follow up CRF.

Participants who become pregnant during the trial will be transferred from the trial to the Digital Surveillance pathway within the screening programme as there are National Guidelines for more frequent than normal screening during pregnancy.

If the participant is identified during follow-up to have developed screen positive retinopathy, significant other eye disease or the images are no longer gradable, the participant will be referred to the Hospital Eye Service for future care. These participants will no longer be followed-up (see 5.3.3.2, C) unless this is proved to be a false positive event in which case they can continue in the trial.

5.3.1 Participant transfers

If a participant moves out of the area, they will be withdrawn from the trial at that point and transferred to another NHS Diabetic Eye Screening Programme where they will be offered routine care for that programme.

5.3.2 Premature discontinuation of trial intervention

If a decision is made by the patient and or/clinician not to continue with personalised screening visits the participant will revert back to routine care. The follow-up of these patients will continue at the next screening interval and data will still be collected unless the participant explicitly withdraws consent for follow-up. Premature discontinuation will be recorded.

5.3.3 Participant withdrawal

Participants are free to withdraw consent at any time without providing a reason. Participants will be provided with the relevant contact details within the participant information leaflet, should they wish to withdraw consent at any time. For participants who withdraw consent prior to randomisation, the data collected up until the point of withdrawal will be used for the analyses.

For patients who withdraw following randomisation the reason and level of withdrawal will be recorded on Form 8: Withdrawal / Lost to follow up.
For this trial routine clinical data will be collected at follow–up visits and at least 700 participants will also be invited to complete Health Economic Questionnaires. The participant can withdraw consent for some or all elements of data collection; these are described in the following sections.

5.3.3.1 Withdrawal from Health Economics

If the participant withdraws from Health Economics data collection only, these patients will continue to attend their follow-up visits and Health Economics data will not be collected at future visits. Health Economics data will be retained unless the patient explicitly withdraws consent for this also.

5.3.3.2 Withdrawal from trial completely

Patients who wish to withdraw consent for the trial will have anonymised data collected up to the point of withdrawn consent and the data (up to the time of withdrawal) will be included in the analyses unless the patient explicitly states that this is not their wish. For instance:

A. Withdrawal from all further follow up and trial data collection - *no further data to be collected, however data collected previously will remain for trial analysis.*
B. Withdrawal from all further follow up and trial data collection, including data previously collected - *no further data will be collected and all data previously collected for the trial will be removed.*

C. Withdrawal from the Data Warehouse - *data contained within the Data Warehouse will be removed.*

5.3.3.3 Process of Managing Withdrawal

If a patient chooses to withdraw consent during a clinic appointment, the researcher will complete a withdrawal form. This will then be received by the Research Clerical Team, who will in turn notify the CTRC.

If a patient informs the Research Clerical Team directly via telephone of their wish to withdraw, the delegated member of the team will complete a Form 8: Withdrawal/Lost to Follow up and enter on to the RCT database and notify the CTRC.

6 ENROLMENT AND RANDOMISATION

6.1 Screening

It is expected that the majority of patients attending for screening for diabetic retinopathy will be eligible for inclusion in the trial.

Pre-screening will be performed by the ISDR research clerical team who will assess eligibility by comparing the screening programme patient database to the ISDR register. The ISDR register contains the details of patients whose GPs have enrolled with the study and who have not opted out of transfer of their data to the Data Warehouse using EMIS Web. Eligible patients will be allocated a unique trial identifier (E-Number). Patient's E-number will be held in a relational database with their NHS number and date of birth. A study invitation letter and a patient information leaflet (PIL) will be posted to the patient approximately one week prior to the screening appointment along with a copy of their screening clinic appointment.

As patients arrive for their appointments a trained researcher will approach these patients to ask them whether they have received, read and understood the information about the study and whether they are interested in participating in the trial. The researcher will also establish that the patient has not opted out of the study Data Warehouse between receiving their invitation letter and attending for screening. (This is unlikely but possible).

Subsequent details of consent processes, including patients who have been consented and reason consent not sought or not provided, will be collated on the Consent and Randomisation Log. Paper CRFs are currently used throughout the clinic for ease of data collection however data will be submitted electronically within 5 days of the end of each clinic to allow review of data at the end of the calendar month.

6.1.1 **Translation**

For those patients who require a translator an appropriate qualified interpreter employed by the Trust will be used at the Clinical Eye Research Centre, Royal Liverpool University Hospital only.

6.2 Consent

Patients will have received a Patient Information Leaflet (PIL), which contains all the details of the trial, with their screening appointment reminder. On attending the screening clinic the researcher will approach the patient to establish whether they wish to participate in the trial. If the patient expresses an interest in participating the researcher will take written informed consent from the patient and formally enrol them into the study. See section 11.3 for further information on informed consent. In order not to interrupt the screening service consent may be taken after the patient has had dilating drops. The researcher will offer to read the consent form to the patient at this point. This will be recorded on the Consent form. This is normal practice when consenting patients in ophthalmology clinics for procedures and treatment.

The researcher will complete the trial Consent and Randomisation Log which records barriers to consent. This will be recorded electronically as stated in section 6.1.

A separate record will be kept of patients who have received study information but not attended for screening. This will be recorded by the Research Clerical Team.

6.3 Baseline

Data for the Baseline Screening Assessment CRF will be collected by the researcher taking patient consent. This will include date of screening, gender, involvement in DR treatment studies and vision data.

Ethnicity, smoking status and whether the patient has completed the Health Economics questionnaire will be collected on the Baseline Screening Assessment CRF and transferred to the data warehouse.

Quality of life and cost data (visit questionnaire) will be collected using a Health Economics questionnaire (see section 8.4.1) on at least the first 700 eligible patients enrolled into the study. Due to time between recruitment, randomisation and communication with the health economics team, the actual number recruited to this sample may exceed 700.

Following the Baseline Screening Assessment, the Baseline Grading Log will be completed providing grading details and outcomes including assessment of the screening images and eligibility for randomisation.

6.4 Eligibility/Enrolment

Confirmation of eligibility can only occur once the patient's digital images have been graded in the central grading centre. The expected timeframe for patients to have their screening result is within 6 weeks, as recommended by the National Screening Programme. For this trial, the 6 week standard is expected to be maintained, however there will be a maximum window of upto 12 weeks to allow for any potential delays in the assessment of digital images, administrative errors and risk factor collection.

The number of patients for whom the receipt of the screening result exceeds the 6 week standard will be monitored. A breakdown of the number of patients for whom this has occurred will be reported. Patients whose results are delayed will be contacted by the ISDR clerical team to explain why the delay has occurred and to confirm that the patient wishes to continue in the ISDR study.

Patients who are not suitable for the trial because they have screened positive DR, other significant eye disease or ungradable digital retinal images will be excluded from the study. They will be informed that they will not be able to continue to take part in the trial in their results letter. They will be referred to the dedicated screen positive assessment clinic at the hospital. This will also be recorded on Consent and Randomisation Log.

All other patients will continue into the trial. For both eligible and non-eligible patients the Consent and Randomisation Log and Baseline Grading assessment will be completed to detail the individual patient outcome. Form 3, Baseline Screened Positive log will be used to collect the LDES and NDESP retinopathy, maculopathy and photocoagulation levels once these patients have been seen in the dedicated assessment clinic.

6.5 Randomisation

The retinopathy grading result will be sent to the ISDR data warehouse as this completes the risk factor data collection for each individual patient

Participants will be randomised using a secure (24-hour) web based randomisation programme controlled centrally by the CTRC. Participants who have been screened will be identified by E-number and this will be used to generate the randomisation number. Block randomisation with stratification will be used to ensure adequate representation in each arm of the trial for groups where there are small numbers of participants (such as children aged 12-15).

Randomisation can only be completed once the patient's images have been graded as it is only possible at this point to complete the inclusion and exclusion criteria. Patients who screen positive due to DR, other eye disease or ungradable images will have been consented but will not be included in the trial. These patients will receive a letter explaining why they have not been included in the trial and their follow up details.

Data on the patient's retinopathy status is fed automatically from the screening software (OptoMize) into the Data Warehouse and then to the Randomisation and RCE programmes. The ISDR data warehouse will be used to automatically populate the majority of the fields in the baseline and follow-up electronic CRFs (OpenClinica). This data will include inclusion and exclusion criteria information, age, and other values of risk variables used to assess risk in the risk engine but not strictly needed for the randomisation. Randomisation and risk allocation is now performed automatically and an automated email confirmation is sent to an authorised researcher and the PI or Co-investigator (where applicable), (initially randomisation was performed manually). Confirmation of the randomisation arm will be sent via email to the Trial Coordinator. An authorised researcher will transfer randomisation details onto the Consent and Randomisation Log.

The risk engine will generate percentage risk and the appropriate screen interval and this information will be fed automatically into the screening software (OptoMize) and also recorded in OpenClinica. OptoMize will generate a letter to the patient informing them of the screening result and their next planned review date. It is a national requirement that patients, and the GP, are informed of the screening result within 6 weeks of their screening appointment see appendix 1.

6.5.1 **Back-Up Randomisation**

In the event of an internet connection failure between the centre and the randomisation system, the centre should contact the CTRC immediately to try to resolve the problem using the backup randomisation system (a mirror of the live system).

Randomisation web access if required:(https://ctrc.liv.ac.uk/Randomisation/ISDR/Home/Login)
If there are any problems with web randomisation, please contact the CTRC
Monday to Wednesday morning : 0151 282 4707
Wednesday afternoon to Friday : 0151 706 4784
Or via email on isdrtrl@liverpool.ac.uk

(Note that the CTRC is open from 0900 – 1700, Monday – Friday, excluding public holidays)

Research staff will be trained to use the web randomisation system during the initiation process. After research staff have been trained they will be issued with personal login and password details.

6.6 Follow up Visits

Once the patient has been randomised follow up data will be collected based on the Randomisation Arm and variable screening interval.

At each follow up appointment patients will have Visual Acuity assessed and images taken. Form 5: Follow up Assessment will record the outcome of the screening assessment.

Patients who complete the Health Economics questionnaire and are subsequently randomised onto the trial will be asked to complete the Quality of Life section at each subsequent visit.

The images taken at the Follow up Assessment will be graded, and assessed for screen positive DR.

After grading, patients randomised to the personalised risk-based screening arm will have their risk re-run and the next screen interval will be determined accordingly.

Patients who are false positive at their first assessment in the Hospital Eye Service, or who are returned immediately to screening (e.g. patients with cholesterol embolus) will be asked if they wish to continue in the trial. If they wish to stay in the trial and they are in the personalised risk based screening arm of the trial their risk will be re-calculated in the RCE and their new interval determined.

7 TRIAL INTERVENTION

7.1 Introduction

Patients will be randomised to two arms, either personalised risk based screening intervals or annual screening (usual care) intervals.

7.1.1 Accountability procedures for study intervention

Failsafe methods in the LDESP will be used to ensure that all patients are informed about their appointments and non-attenders are followed up. This is an established part of the LDESP. Participants who fail to attend their screening appointment will be given a further appointment, usually within 6 weeks. If they fail to keep the 2nd appointment, a letter will be sent to them and their GP asking them whether they wish to opt-out of the screening programme or re-book a further appointment. If they do not respond, then they will be recalled again in the next round of screening. Patients in the 6 month personalised risk based screening arm of the trial who fail to attend their screening appointment on two occasions and who do not wish to opt out of screening will be offered a new screening appointment in 6 months' time. Patients in the fixed (annual) arm of the trial and in the 12 or 24 month personalised risk based screening arm of the trial who fail to attend their screening appointment on two occasions and who do not wish to opt out of screening will be offered a new screening appointment in 12 months' time.
If they choose to opt-out of the screening programme, they can do so for up to 3 years and they will not receive another screening appointment during that time. If they choose to re-book a further appointment and they fail to keep that one, then a further letter will be sent to them and their GP asking them whether they wish to opt-out of the screening programme or re-book a further appointment. This cycle of re-booking appointments and failure to attend can continue indefinitely until either the participant indicates that they wish to opt-out of the screening programme or they fail to respond to a letter. For the purposes of this study, we will consider anyone who has not attended a screening appointment within 90 days of the expected follow-up date to have not attended that particular visit. However, they will still follow the standard screening procedure as described above.

7.1.2 Assessment of compliance with study intervention

A comparison of compliance (attendance) is the primary objective of the study. See section 7.1.1.Compliance data will be transferred to the CRF.

7.2 Administration of trial intervention

The randomisation arm will be allocated once eligibility has been established – see Section 6.

7.3 Concomitant Medications/Treatments

7.3.1 Medications permitted

All medications are permitted.

7.3.2 Data on concomitant medication

Data will be collected through patients' own GP medical records on medications prescribed for diabetes, high blood pressure and high cholesterol.

7.4 Co-enrolment guidelines

Patients who enter any trial investigating a treatment aiming at preventing or modifying the development of STDR will exit the ISDR trial.

8 ASSESSMENTS AND PROCEDURES

8.1 Schedule for follow-up

The trial will continue for 54 months (19 months for recruitment) to ensure that all patients in the extended screen interval group (24 months) have attended for follow up screening at least once (24 months follow up plus 90 day window). For the trial, a 90-day window will be considered from the time when an expected follow-up is due. If an individual does not attend a screening appointment within this window then they are classed as not attending an appointment at that specified follow-up. However, in practice, if patients do not attend their first appointment for screening they are offered a second appointment (usually within 6 weeks of the first appointment). If patients fail to attend for their second appointment a letter is sent to the patient advising them of the importance of screening and asking them to either contact the screening programme for a new appointment or request to opt-out of the screening programme altogether. If an individual opts out of the screening programme altogether, then they will not receive another screening appointment for up to 3 years. If an individual does not respond to the letter then they will be recalled to screening at the next scheduled visit (i.e. in 6, 12 or 24 months' time from the original appointment date). If they request a further appointment and fail to attend, the same letter will be sent to the individual and their GP, with the same options. This cycle will continue indefinitely until either the participant decides to opt-out of the screening programme or does not respond.

				Follow-Up Schedule							
Procedures	*Screening*	*Baseline*	*randomisation*	*6 months***	*12 months***	*18 months***	*24 months***	*30 months***	*36 months***	*Study Completion*	*Premature Discontinuation*
Signed Consent Form	X										
Visual acuity(LogMar)		X		X	X	X	X	X	X		
Visit Questionnaire		X									
QoL Questionaire (HUI3 and EQ5D)*		X		X	X	X	X	X	X		
Assessment of Eligibility Criteria	X	X	X								
Randomisation			X								
Digital imaging		X		X	X	X	X	X	X		
Acceptability questionnaire									X		

*The Health Economics questionnaire, incorporating the Quality of Life (QoL) questionnaires EQ 5D-5L [37, 38] and Health Utilities Index Mark 3 [HUI3] [38] and the Visit questionnaire, will be completed on approximately 700 eligible patients enrolled into the study. The Visit questionnaire will not be repeated at follow up visits. The QoL questionnaires will be repeated at every subsequent follow up visit for those patients who completed it at baseline. (See Section 8.4.1).

**Patients will be randomised to 2 arms. In one arm patients will only attend at 12 month intervals. In the other arm patients will attend at 6, 12 or 24 monthly intervals. Note that all patients are expected to be followed up for at least 24 months. Those patients recruited

during the first months of the trial may be followed up for a longer period than 24 months (30 months and/or 36 months) and the corresponding data will still be collected.

8.2 Procedures for assessing efficacy

Data required to record activity / efficacy is as outlined above and in the secondary outcomes in Section 4.2. The following items will be mainly collected retrospectively by the trial administrative team and the research nurses:

- Development of STDR
- Retinopathy level at screening (LDES and NDESP grading)
- Maculopathy level at screening (LDES and NDESP grading)
- Date of screening appointment/s.
- Attendance at screening appointment/s
- Referral to dedicated diabetes assessment clinic
- Visual acuity (logMAR) at dedicated diabetes clinic appointment
- Outcome of dedicated diabetes assessment clinic
- Attendance at follow up eye clinic
- Date of treatment/s (laser / intraocular injection) for STDR
- Date of surgery for STDR
- New visual impairment (using the following two thresholds: ≥ +0.50 logMAR and ≥ +0.30 logMAR in the best eye)
- New visual impairment due to diabetic retinopathy (using the following two thresholds: ≥ +0.50 logMAR and ≥ +0.30 logMAR in the best eye)

8.3 Procedures for assessing safety

Safety data will be assessed by the Independent Safety and Data Monitoring Committee (IDSMC) for the trial. The IDSMC will be embedded within the Programme Steering Committee (PSC). All members of the PSC are independent and have signed Conflict of Interest documents. The IDSMC will sit at least annually. Day to day monitoring of the trial will be under the remit of the Trial Management Group (TMG) which will meet at least 3 monthly and monthly in the first 6 months of the trial. Safety measures will be assessed at each meeting and highlighted to the IDSMC if concerns are raised.

Potential risks are described in Section 2.8.1.

The following safety parameters to be collected:

- Withdrawal of consent rates in both arms of the trial
- Attendance rates in both arms of the trial
- STDR detection rates in both arms of the trial
- New visual impairment due to diabetic retinopathy in both arms of the trial.
- Treatment rates in both arms of the trial
- Hospital attendance for diabetic life-threatening events (heart attack and stroke) in both arms of the trial
- Attendance at GP annual review rates in both arms of the trial

Measures will be put in place in the Hospital Eye Service for identification and reporting of patients presenting symptomatically with STDR.

8.4 Other assessments

8.4.1 Quality of life and health economics

Patients enrolled into the study will be asked to complete the Health Economics questionnaire (including QoL and Visit questionnaires) once they have consented to enter the trial (see section 6.3). These will be self-completed by the participant where possible. A large print paper version will be available and individuals will receive assistance from the researcher wherever required. If assistance is required this will be recorded and the researcher will adhere to a specific script to ensure consistency. There will be two types of questionnaire. The Visit questionnaire will be about the individual's travel, time foregone and any personal expenses associated with the visit. The QoL questionnaire will include the EQ-5D-5L and Health Utilities Index Mark 3 (HUI3) questionnaires. The two questionnaires should take around 15 minutes to complete. The randomisation programme will inform the programme when approximately 700 patients have completed the Health Economics questionnaire have been enrolled into the study.

The QoL questionnaire will be repeated each time these participants attend for screening. This should only take around 8 minutes at each subsequent visit.

If a participant has consented to the trial but not been enrolled because they have confirmed STDR they may be sent a copy of the health questionnaire to complete in their own time at home in the future. Similarly, participants who subsequently develop STDR during the trial may also be sent a copy of the health questionnaire to complete at home.

In addition to trial outcomes, the data collected using these questionnaires will be used to inform the modelling work carried out as a part of Workstream D. This will require the interim analysis of baseline QoL and Visit Questionnaire data. These data will be analysed independent of other trial data (including trial allocation) in order to estimate the quality of life of people attending screening and to estimate the costs associated with attending screening.

The researcher will transcribe the data from the Visit and QoL Questionnaires onto the electronic CRF.

8.5 Sub studies

8.5.1 Workstream F Questionnaire

Workstream F (REC approval NRES Committee North West – Preston
28 March 2013 Ref 13/NW/0196) comprises two phases;
Phase I involves semi-structured qualitative interviews with patients who regularly attend eye screening, and with a range of health professionals involved in the delivery of eye screening services. We explore a range of issues surrounding the acceptability of moving from annual to risk-based, extended screening intervals.
In phase II will involve semi-structured qualitative interviews with patients involved in the variable screening arm of the RCT. We will explore their experiences and views of moving to variable intervals.

Once the questionnaire has been developed it will be submitted as an amendment to both protocols.

8.5.2 **MRC START**

ISDR will be acting as a host trial to test an MRC (Medical Research Council) START (Systematic Techniques to Assist Recruitment to Trials) recruitment intervention. The embedded trial (MRC START in ISDR) is a collaboration with the University of Manchester. Part of a programme of work funded by the MRC methodology research programme, ISDR would become one of 11 similar trials conducting embedded studies of recruitment interventions that are designed to build the evidence base around RCT methodology.

The proposed embedded trial would test whether patient information materials, including consent forms, which have undergone an optimisation process, improve recruitment to the ISDR study. The ISDR recruitment procedures would remain unchanged, but patients eligible for ISDR would be randomised to receive the standard or optimised versions of the patient information materials and the outcome (recruitment to ISDR) would be recorded.

8.6 Loss to follow-up

For patients who DNA their screening appointment, loss to follow up will be managed as part of the existing failsafe procedures with LDESP (see section 7.1.1). However, following randomisation and throughout the study there may be instances where the patient will be considered as lost to follow up. Examples can be found below:

- Death
- Patient wishes to discontinue trial intervention (will revert to standard intervention and will continue to be followed up)
- Moved to a non-participating GP practice
- Screen positive diabetic retinopathy at follow up screening
- Diagnosed with screen positive significant other eye disease requiring referral to HES at follow up assessment(s).
- Log MAR/ VA Unobtainable at follow up Screening assessment(s)
- Patient refused to have visual acuity (Log MAR) assessed at follow up screening assessment(s)
- 8 =Unobtainable / ungradable images at follow up assessment

If a patient loses the capacity to consent during the trial all data collected up to the point they lost capacity will be included in the trial but no further data will be collected. If the patient is considered lost to follow up **before randomisation**, all reasons should be recorded on the Consent and Randomisation log. If the patient is considered loss to follow up **after randomisation**, this is recorded on Form 8: Withdrawal and Lost to Follow up.

The Researcher must document the reason(s) on the appropriate section of the RCT database. Information can be collated using a paper CRF, however should be transferred to the RCT database ideally within 24 hours of becoming aware of the loss to follow up.

8.7 Trial closure

The end of the trial is defined to be the date on which data for all participants is frozen and data entry privileges are withdrawn from the trial database. However, the trial may be closed prematurely by the Trial Steering Committee (TSC), on the recommendation of the Independent Data and Safety Monitoring Committee (IDSMC).

9 STATISTICAL CONSIDERATIONS

9.1 Introduction

A separate and full Statistical Analysis Plan will be developed prior to the final analysis of the trial. The statistical analysis plan will be agreed by the Trial Steering Committee before being sent to the IDSMC for comments and approval. The main features of these planned statistical analyses are included in this protocol.

9.2 Method of randomisation

Patients will be randomly allocated to the personalised risk-based screening interval or the usual fixed annual interval once inclusion criteria are confirmed (for example, patients who give consent to participate in the trial but whose photographic images some weeks later indicate STDR, will be excluded from the trial).

Randomisation will be carried out at individual level using a web-based randomisation system by the trial administrative team. Block randomisation with stratification will be used to ensure adequate representation in each arm of the trial for groups where there are small numbers of participants (Patients will be stratified by clinic and age; <16 years old and ≥16 years old). Patients are being recruited from all the screening centres in Liverpool. These have different proportions of patients in hard to reach groups, such as minority ethnic groups and degree of social deprivation. We believe therefore that the patients recruited will be representative of the screening populations as a whole.

Once a patient is randomised and the patient has been allocated to the personalised risk-based screening interval arm the risk engine will be used to generate the individual screening interval using the patient's retinopathy grading and other risk factors.

Randomisation can only be completed once the patient's images have been graded. Because of this there are a number of weeks from consent to confirmation of eligibility and randomisation. Due to this delay, it is expected that the final number of participants recruited is likely to be higher than that which is specified in section 9.4.2. Given the rate of recruitment, it is not anticipated that the recruitment target will be exceeded by anymore than 100 participants above the specified recruitment target.

9.3 Outcome measures

See Section 4

9.4 Sample size

Original and revised sample size calculations are included. Sample size revisions are necessary due to lower patient recruitment rates than expected.

Data from the Liverpool retinal screening programme suggests an attendance rate for annual follow-up screening of around 85%. However, nationally, estimates of attendance rates are closer to 75%. Data from the existing programme indicates that 13,924 patients attended for screening in 2010/11. With numbers increasing slightly each year we can estimate a screen population of 14,000. All 93 GP practices now have EMIS web. Assuming about 76% of eligible GPs participate and a conservative 70% consent rate, we expect about 6900 patients enrolling. After deducting 3% due to STDR at baseline, we expect approximately 6700 patients available for the RCT. The annual rate of loss to follow-up due to death or

exclusion from screening (physical or mental bar to screening, patient opt-out) is around 6%. This would leave a sample of around 5900 available for follow-up at 2 years.

9.4.1 **Original sample size justification**

Since the number of patients available per arm over 12 months is expected to be enough to analyse differences in attendance rates between the two pathways, we also consider the sample size required to show non-inferiority for detection of STDR in the RCT (secondary question).

The STDR detection rate predicted for the usual care pathway during the two years follow-up is approximately 6%, based on data from the LDES. Applying the sample size equation developed by Machin and Campbell (see Equation 5.7) [39] approximately 1,950 patients per arm are required (a total of 3,900) to detect a difference ≥5% in attendance rate between the two pathways. This is based on an equivalence study with power 95%, significance level 5% and equivalence of 0.05 where an attendance rate of 75% has been assumed for the usual care pathway. Therefore, the number of patients available (5900) will permit the detection of a maximum allowable reduction of 1.5% in STDR detection of the individualised care pathway compared to standard care, with 5% significance level and 75% power.

In practice we will recruit as many participants as we can within the recruitment period, to enable us to collect enough data to properly assess other outcomes, particularly safety outcomes such as the number of people attending screening with retinopathy levels already above the normal referral rate.

9.4.2 **Revised sample size justification**

Due to a lower recruitment rate and lower attendance rate than expected the sample size is revisited. Additionally, adjusting the power required from 95% to 90%, the number of patients required to address the primary question and the minimum required number of patients is 3940.

Taking into account a loss to follow-up rate of 6% due to death and other exclusion from screening (note that non-attendance is the primary outcome and therefore it is not factored here), the target for recruitment (patients to be randomised into the trial) is 4460 patients. We expect this number will allow us to retain sufficient patients (3940 patients) for the primary analysis since this is the number of patients retained after applying a 6% loss to follow-up rate per year 4460*0.94*0.94=3940).

We have also reviewed the sample size required to show non-inferiority for detection of STDR in the RCT. The STDR detection rate predicted for the usual care pathway during the two years follow-up is approximately 6%, based on data from the LDES. The sample size required (n=4460 patients randomised with 3940 patients retained after 2 years from baseline) will permit the detection of a maximum allowable reduction of 1.5% in STDR detection of the individualised care pathway compared to standard care, with 5% significance level and power between 60-65%.

9.5 Interim monitoring and analyses

Formal interim analyses of the accumulating data will be performed at regular intervals (at least annually) for review by an IDSMC. These analyses will be performed at the CTRC. The IDSMC will be asked to give advice on whether the accumulated data from the trial, together with results from other relevant trials, justifies continuing recruitment of further patients or further follow-up. A decision to discontinue recruitment, in all patients or in selected subgroups will be made only if the result is likely to convince a broad range of clinicians including participants in the trial. If a decision is made to continue, the IDSMC will advise on the frequency of future reviews of the data on the basis of accrual and event rates. The IDSMC will make recommendations to the Trial Steering Committee (TSC, see section 16) as to the continuation of the trial.

9.6 Analysis plan

A full and detailed statistical analysis plan will be developed prior to the final analysis of the trial. The main features of the statistical analysis plan are included here.

The main aim will be to assess the relative safety of personalised risk based interval screening as measured by attendance rate. We will test the hypothesis that the difference in attendance rates between the two pathways is within an acceptable range by assessing the bounds of the 90% confidence interval for the difference in attendance rate (5% significance level, 95% power, acceptable range of 0.05).The relative safety of personalised risk based interval screening to detect STDR will also be assessed (non-inferiority analysis).

We will undertake sub-group analyses to assess differences in attendance rates between the two arms for three different risk groups (the risk groups will be defined based on the individual baseline estimated risk of being screen positive). Additionally, we will investigate demographic and clinical factors related to non-attendance. A logistic mixed effects model that takes into account the patient's covariate information at baseline (HbA1c, systolic blood pressure, total cholesterol, disease duration, retinopathy level, ethnicity, age, gender, smoking status, diabetes type) and screening clinic (clustered data) will be fitted with attendance (Yes/No) at the first screening visit as the outcome variable. The random effects of the model will account for the variability by screening clinic.

We will apply sensitivity analyses to check the sensitivity of the results on the assumption that missing data are missing completely at random. We will consider simple imputation using the most and least conservative scenarios, as well as multiple imputation.

A Health Economics (HE) Analysis Plan (HEAP) that focusses on the trial-based economic evaluations has been developed. A separate HE report will also be produced. Using trial data and information collected in WS D the cost-effectiveness of variable screening will be evaluated. Bootstrapped ICERs for secondary outcomes and cost-effectiveness acceptability curves for variable versus standard screening will be presented.

10 SAFETY REPORTING

In the ISDR trial there is no treatment intervention. Randomised participants will receive the same type of screening for diabetic retinopathy with the only difference being the interval of screening.

At the patient's subsequent visit for screening at their allocated interval the risk of diabetic retinopathy will again be assessed for patients in the variable screening arm so that, should risk increase, participants are followed up appropriately.

The Randomisation CRF will capture the percentage of risk and the allocated screen interval. The same information will be recorded on the Follow Up Grading Log CRF. This will also be monitored as part of the monitoring plan.

11 ETHICAL CONSIDERATIONS

11.1 Ethical considerations

The study will abide by the principles of the World Medical Association Declaration of Helsinki (1964) and the Tokyo (1975), Venice (1983), Hong Kong (1989) and South Africa (1996).

We do not believe that there are major ethical issues in the trial beyond the normal issues of informed consent and options to withdraw without affecting medical care. There are no extra visits involved in the trial.

Vulnerable populations
We will be recruiting children over the age of 12 in line with current guidance issued by the NDESP. Our Programme Development Study showed that this was acceptable and effective.

Additional tests
Questionnaires will be administered during the first visit taking an estimated 15 minutes to complete. At subsequent visits the questionnaire should only take around 8 minutes to complete.

Risks of participating in the trial
We believe that these are minimal. There is a possibility that STDR may be missed if patients do not attend for routine follow-up. This will be minimised by providing information on the importance of attendance at screening in the Patient Information Sheets.

11.2 Ethical approval

Preston Ethics Committee has given a favourable opinion on the protocols for Workstreams B, C and F of the ISDR programme grant and will be approached to provide Ethical approval for this Workstream (E).

Consent from the patient or legally acceptable representative will be obtained prior to participation in the trial, after a full explanation has been given of the options, including the conventional and generally accepted interval for screening. Patient Information Leaflets (PIL) and Informed Consent Forms will be implemented. The right of the patient to refuse to give consent to participate in the trial without giving reasons will be respected. After the patient has entered the trial, the clinician will remain free to amend the screen interval at any stage, if he/she feels it to be in the best interest of the patient. However, the reason for doing so will be recorded and the patient will remain within the trial for the purpose of follow-up and data analysis according to the option to which they have been allocated. Similarly, the patient will remain free to withdraw at any time from the intervention and trial follow-up without giving reasons and without prejudicing further treatment.

Proxy consent from the parent or legally acceptable representative will be obtained prior to each patient participating in the trial, after a full explanation has been given of the options, including the conventional and generally accepted interval for screening. Age and stage-of development specific PIL and Consent forms will be implemented and patient assent will be obtained where appropriate (see section11.3.1). The right of the parent/ legal representative to refuse consent for the minor to participate in the trial without giving reasons will be respected. After the patient has entered the trial, the clinician will remain free to amend the screen interval, at any stage, if he/she feels it to be in the best interest of the patient. However, the reason for doing so will be recorded and the patient will remain within the trial for the purpose of follow-up and data analysis. Similarly, the parent/legal representative of the patient remains free to withdraw the patient at any time from the protocol intervention

and trial follow-up without giving reasons and without prejudicing the further treatment of the minor.

11.3 Informed consent process

Informed consent is a process initiated prior to an individual agreeing to participate in a trial and continues throughout the individual's participation. Informed consent is required for all patients participating in CTRC coordinated trials. In obtaining and documenting informed consent, the investigator should comply with applicable regulatory requirements and should adhere to Good Clinical Practice (GCP) and to the ethical principles that have their origin in the Declaration of Helsinki.

Participants will be sent an information leaflet approved by an independent ethical committee with their screening appointment 4-6 weeks prior to their screening visit. Consent will be obtained at the screening appointment provided the patient (parent or legal representative in the case of minors) is able to give informed consent at this time. Discussion of objectives, risks and inconveniences of the trial and the conditions under which it is to be conducted will be provided to patients by research nurses with experience in obtaining informed consent. If the patient (parent or legal representative in the case of minors) is not able to give consent they will continue under current care.

The information will emphasise that participation in the trial is voluntary and that the participant may withdraw from the trial at any time and for any reason. All participants will be given the opportunity to ask any questions that may arise. A contact point where further information about the trial may be obtained will be provided.

The participant (parent or legal representative in the case of minors) will sign and date the informed consent document. Both the person taking consent and the participant must personally sign and date the form. A copy of the informed consent document will be given to the participant and their legally acceptable representative for their records. The original copy will be filed in the Investigator Site File. One final copy of the consent form will be sent to the ISDR Research Office.

The participant will be asked to sign the following consent forms as appropriate:

- Consent form for minors (as defined as participants under the age of 16)
- Consent form for adults (as defined as participants over the age of 16)

The participant may, without being subject to any resulting detriment, withdraw from the trial at any time by revoking the informed consent. Similarly, the parent or legal representative may withdraw a minor under the same conditions. The rights and welfare of the participants will be protected by emphasising to them that the quality of medical care will not be adversely affected if they decline to participate in this study.

11.3.1 **Assent in minors**

Young people over 16 years will give informed consent, as described in section 11.3. In line with best practice, young people under the age of 16, if capable and under appropriate circumstances, should be approached to provide assent, by a member of the research team. Proxy consent of parents/guardians/legal representatives of young people under 16 will also be obtained. The process outlined in the previous section 11.3 will apply for parents/guardians/legal representatives of the young person or child.

All information sheets and assent forms will be made available prior to the visit, in age-and-stage-of-development appropriate formats and will include a contact point where further information about the study can be obtained prior to the assent/consent process, the researcher will talk through the outline description including the aims and objectives and specific details of the study as they apply to them will be explained and any questions answered.

The minor should personally write their name and date the assent form, which will then be signed by the parent/guardian/legal representative and the researcher.

Assent forms do not substitute for the consent form signed by the patient's legally acceptable representative. Assent should be taken where appropriate and documented in the patient notes; however the absence of assent does not exclude the patient, provided consent has been obtained from the parent/guardian/legal representative, from participation in the trial.

A copy of the signed assent documentation will be given to the young person and their parent/guardian/legal representative for their records; the original copy will be filed in the participant's medical notes, and a copy will be kept in the research file held by the delegated researchers.

11.3.2 **Re- Consent of 16 year olds**

A participant involved in the trial who reaches the age of 16 (and is therefore deemed competent to provide consent) should be re-consented at their next screening appointment if the researcher is available to take consent.

11.4 Study discontinuation

In the event that the study is discontinued, patients will revert to NDESP guidelines for screen intervals.

12 REGULATORY APPROVAL

This trial does not fall within the remit of the EU Directive 2001/20/EC, transposed into UK law as the UK Statutory Instrument 2004 No 1031: Medicines for Human Use (Clinical Trials) Regulations 2004 as amended. This trial has not been registered with the MHRA and does not require a Clinical Trial Authorisation (CTA). It will be registered International Standard Randomised Controlled Trial Register.

13 TRIAL MONITORING

Trial monitoring will be carried out to ensure that the rights and well-being of human participants are protected during the course of a clinical trial. A risk assessment will be performed for each trial coordinated by the CTRC to determine the level and type of monitoring required for specific hazards. The nature and extent of monitoring will be specific to the individual trial.

Trial Oversight Committees related to the monitoring of the trial are detailed in section 16

13.1 Risk assessment

A risk assessment will be performed prior to initiation to assess the level of risk in accordance with the CTRC SOP TM005.

The risk assessment will be completed in partnership between:

- Representative/s of the Trial Sponsor
- Chief Investigator
- Trial Coordinator and supervising Trial Manager
- Trial Statistician and supervising Statistician
- Information Systems team
- CTRC Director

In conducting this risk assessment, the contributors that will be considered are potential patient, organisational and study hazards, the likelihood of their occurrence and resulting impact should they occur.

The outcome of the risk assessment is expressed as a percentage, assigned according to the following categories:

- Score ≤ 33% = Low risk
- Score ≥34 to ≤ 67% = Moderate risk
- Score ≥ 68 to ≤ 100% = High risk

13.2 Source documents

The CRF will be considered the source document for the following data:

- Date of conducting informed consent/assent
- Patient Ethnicity and Smoking Status
- Attendance at clinic following randomisation
- Participation in Health Economics (QOL and Visit questionnaires)

Data will be added to the study database from the following sources:

- Study number (E- number) – relational database
- Randomisation number – web-based randomisation programme
- Randomisation arm – web-based randomisation programme
- Gender – OptoMize database
- LogMAR vision – OptoMize database
- Attendance rates – OptoMize database

- STDR detection rates – OptoMize database

The following data will also be collected and used in the analyses and reporting of outcomes not collected Open Clinica:

- New visual impairment due to diabetic retinopathy – PENS electronic patient record
- Treatment rates – IPM via Quality Department, RLUH
- Hospital admission for diabetic life-threatening events (heart attack and stroke) – IPM via Quality Department, RLH
- Attendance at GP annual review – EMIS Web

The screen interval for patients in the personalised risk based screening intervals arm of the trial will be determined by the risk engine using the trial dataset derived from data stored in the data warehouse held at the Royal Liverpool University Trust.

In addition the fact that the patient is participating in a clinical trial, date of provision of patient information, date of conducting informed consent/assent, study number and randomisation arm will be added to the patient's medical record (where they exist) chronologically.

13.3 Data capture methods

13.3.1 **Case report forms**

The ISDR study will use a mixed method approach to data collection. The majority of the data will be imported from the Data Warehouse; however paper CRFs will be used during clinical visits to capture specific clinical data items key as well as the participant completed Health Economics surveys. Data entry of the paper CRF items should occur after patient visits (or as close as possible) to allow formal review of data.

For data imported from the Data Warehouse, data will be transferred at regular intervals from the ISDR data warehouse, held at Royal Liverpool and Broadgreen University Hospital Trust.

All data for the trial will be stored electronically into a study specific database held in CTRC.

13.4 Central Monitoring

For data completed via paper CRFs, data will be checked for missing or unusual values (range checks) and checked for consistency with participants over time. Any suspect data will be highlighted in the form of data queries.

Data queries will be produced from the trial database. Data queries will be allocated to a named individual (as listed on the site delegation log) usually the researcher. Data queries can be viewed on the trial database and sent electronically or a list of data queries will be provided by the Data Manager to a named individual (as listed on the site delegation log). The delegated individual will respond to the queries providing an explanation/resolution of the discrepancies and will submit the data query for review. The data query will then be stored electronically along with the appropriate electronic data and the appropriate corrections made on the database. There are a number of monitoring features in place at the CTRC to ensure reliability and validity of the trial data, to be detailed in the trial monitoring plan.

As part of the safety reporting data will be captured on the Baseline Grading log and Follow-up Grading Log CRFs and detailed in the monitoring plan.

13.5 Clinical site monitoring

In order to perform their role effectively, the trial coordinator (or monitor) and persons involved in Quality Assurance and Inspection may need direct access to primary data, e.g. patient records, laboratory reports, appointment books, etc. Since this affects the patient's confidentiality, this fact is included on the Parent Information Leaflet (PIL) and Informed Consent Form.

13.5.1 **Confidentiality**

Individual participant medical information obtained as a result of this study is considered confidential and disclosure to third parties is prohibited with the exceptions noted below. CRFs will be labelled with the patient's DOB, initials and unique randomisation number. Medical information may be given to the participant's medical team and all appropriate medical personnel responsible for the participant's welfare.

The CTRC will be undertaking activities requiring the transfer of identifiable data. Verification that appropriate informed consent is obtained will be enabled by the provision of copies of participant's signed informed consent/assent forms being supplied to the CTRC by recruiting centres, which requires that name data will be transferred to the CTRC.

This transfer of identifiable data is disclosed in the PIL and Consent forms. The CTRC will preserve the confidentiality of participants taking part in the study and The University of Liverpool is registered as a Data Controller with the Information Commissioners Office.

13.5.2 **Quality assurance and control**

Quality assurance (QA) includes all the planned and systematic actions established to ensure the trial is performed and data generated, documented/recorded and reported in compliance with applicable regulatory requirements. Quality control (QC) includes the operational techniques and activities done within the QA system to verify that the requirements for quality of the trial-related activities are fulfilled. To this end:

- The research team will attend a launch meeting co-ordinated by the CTRC in conjunction with the Chief Investigator, which will incorporate elements of trial specific training necessary to fulfil the requirements of the protocol.
- The Trial Co-ordinator is to verify appropriate approvals are in place prior to initiation of a centre and the relevant personnel have attended trial specific training.
- The Data Manager is to conduct data entry consistency checks and follow-up data queries.
- Independent oversight of the trial will be provided by the IDSMC and independent members of the TSC.

13.6 Records retention

The investigator at each investigational site must make arrangements to store the essential trial documents (as defined in Essential Documents for the Conduct of a Clinical Trial (ICH E6, Guideline for Good Clinical Practice)), including the Investigator Site File, until the CTRC informs the investigator that the documents are no longer to be retained, or for a maximum period of 15 years (whichever is soonest).

In addition, the investigator is responsible for archiving all relevant source documents so that the trial data can be compared against source data after completion of the trial (e.g. in case of inspection from authorities).

The investigator is required to ensure the continued storage of the documents, even if the investigator, for example, leaves the clinic/practice or retires before the end of required storage period. Delegation must be documented in writing.

The CTRC will archive documents in compliance with ICH GCP utilising the Records Management Service of the University of Liverpool. All electronic CRFs and trial data will be archived onto an appropriate media for long term accessible storage. Hard copies of data will be boxed and transferred to specially renovated, secure, premises where unique reference numbers are applied to enable confidentiality, tracking and retrieval.

14 INDEMNITY

ISDR – Workstream E is sponsored by the Royal Liverpool University Hospital and the University of Liverpool and co-ordinated by the CTRC in the University of Liverpool. The Royal Liverpool University Hospital and the University of Liverpool do not hold insurance against claims for compensation for injury caused by participation in a clinical trial and they cannot offer any indemnity. As this is an investigator-initiated study, The Association of the British Pharmaceutical Industry (ABPI) guidelines for patient compensation by the pharmaceutical industry do not apply. However, in terms of liability, NHS Trust and Non-Trust Hospitals have a duty of care to patients treated, whether or not the patient is taking part in a clinical trial, and they are legally liable for the negligent acts and omission of their employees. Compensation is therefore available in the event of clinical negligence being proven.

Clinical negligence is defined as:
"A breach of duty of care by members of the health care professions employed by NHS bodies or by others consequent on decisions or judgments made by members of those professions acting in their professional capacity in the course of their employment, and which are admitted as negligent by the employer or are determined as such through the legal process".

15 FINANCIAL ARRANGEMENTS

This trial is part of a programme grant funded by the National Institute for Health Research (NIHR). As the contractor, the Royal Liverpool and Broadgreen University Hospital Trust (RLBUHT) receive quarterly payment from the funder which is then held in a designated control code account.

Research collaboration agreements are in place between the main contractor RLBUHT and external collaborators: the University of Liverpool, the University of Nottingham and Gloucester Hospitals NHS Foundation Trust. Invoices of expenditure generated by collaborators will be ratified and paid quarterly by the RLBUHT.

This study will be automatically adopted onto the NIHR Clinical Research Network Portfolio as a programme grant funded by the NIHR. It is eligible to receive research funding support from the local comprehensive local research network (Cheshire and Merseyside) if required.

Trial participants will not be paid to participate in the trial.

16 TRIAL COMMITTEES

16.1 Trial Management Group (TMG)

A Trial Management Group (TMG) will be formed comprising the Chief Investigator, other lead investigators (clinical and non-clinical), experts in the field of screening and members of the CTRC Clinical Trial Unit. The TMG will be responsible for the day-to-day running and management of the trial and will meet at least 3 monthly and monthly for the first 6 months of the trial. Refer to the TMG terms of reference and trial oversight committee membership document for further details.

16.2 Trial Steering Committee (TSC)

The Trial Steering Committee will consist of an independent chairperson, independent clinicians and statisticians with relevant expertise, at least one individual who is able to contribute a patient and/ or wider public perspective. Appropriate members of the TMG will be invited including a sponsor representative and a representative from the research network. The TSC will have a minimum of 75% majority of independent members. The role of the TSC is to provide overall supervision for the trial and provide advice through its independent Chairman. The ultimate decision for the continuation of the trial lies with the TSC. Refer to the TSC terms of reference and trial oversight committee membership document for further details.

16.3 Independent Data and Safety Monitoring Committee (IDSMC)

The Independent Data and Safety Monitoring Committee (IDSMC) will consist of an independent chairperson, plus 1 independent member who is an expert in the field of diabetic retinopathy and a biostatistician.

The IDSMC will be responsible for reviewing and assessing recruitment, interim monitoring of safety and effectiveness, trial conduct and external data. The IDSMC will first convene prior to initiation and will then define frequency of subsequent meetings (at least annually). Details of the interim analysis and monitoring are provided in section 9.

The IDSMC will provide a recommendation to the Trial Steering Committee concerning the continuation of the study. Refer to the IDSMC charter and trial oversight committee membership document for further details.

17 PUBLICATION

The results will be analysed and published as soon as possible but only with the prior consent of the Trial Management Group.

The Trial Management Group will form the basis of the Writing Committee and advise on the nature of publications. The Uniform Requirements for Manuscripts Submitted to Biomedical Journals (http://www.icmje.org/) will be respected. All publications will include a list of participants, and if there are named authors, these will include the Programme Grant's Chief Investigator, the trial Principal Investigator, Statisticians and Trial Managers involved at least. If there are no named authors (i.e. group authorship) then a writing committee will be identified that would usually include these people, at least. The ISRCTN allocated to this trial will be attached to any publications resulting from this trial.

Publications will carry the NIHR acknowledgement and disclaimer. The disclaimer will be required on all outputs from the programme grant, and the full acknowledgement will be added to any output containing research findings.

The members of the TSC and IDSMC will be listed with their affiliations in the Acknowledgements/Appendix of the main publication.

18 PROTOCOL AMENDMENTS

18.1 Version 1 (08/Jan/2014)

Original Approved version.

18.2 Version 2 (28/May/2014)

List of revisions to Version 1 of the Protocol now incorporated in Version 2:

Section	Page	Amendments
N/A	1	ISRCTN added to page 1.
8.4.1.	28	Additional information regarding participant involvement in the QoL Questionnaire.
9		Section on sample size considerations updated. The required sample size is now calculated to only focus on the primary question (attendance rates). Comments made by Chris Rogers addressed.

18.3 Version 3 (03/November/2015)

List of revisions to Version 2 of the Protocol now incorporated in Version 3:

Section	Page	Amendments
N/A	N/A	Clarifications and small amendments have been made throughout the Protocol.
N/A	3	Addition of lead statistician authorisation.
N/A	6	Removal of Medical Expert details, as safety reporting is not applicable.
1	11	Addition of recruitment target to summary, clarification of study duration (same changes to section 8.1).
2.4 to 2.6	15-17	Minor corrections and further information to supporting background information.
5.3	21	Additional Information on participants who become pregnant.
5.3.2	22	Sections removed and Levels of withdrawal described in further detail, see subsequent sections.
6.1 and 6.1.1	23	Clarifications, further detail and minor wording changes for screening process
6.2	23	Further detail provided regarding the consenting process at site, removal of last two paragraphs information to be included in 6.4.
6.4	24	Sequence of information r
6.5	25	Randomisation process amended and clarified to include detail regarding its automated status. Contact details amended.
6.6	26	Further information and wording clarifications
8.4	29/30	Further detail on health economic data collection and analysis.
8.5.2	30/31	Additional sub-study, MRC START.
8.6	31	Loss to follow-up section re-written and extended.
9.4	32/33	Amended Sample size and re-calculations to reflect revised recruitment target. Addition of justifications.
9.6	34	Further detail provided
13.2	40/41	Amendments/ additions to source documentation list
13.3 and	41	Further detail and clarification on data capture methods and monitoring.

13.4		
16.2	46	Change in TSC membership conditions
17	47	Additional publication information.

18.4 Version 4 (11/May/2016)

List of revisions to Version 3 of the Protocol now incorporated in Version 4:

Section	Page	Amendments
N/A	N/A	Clarifications and small amendments have been made throughout the Protocol.
N/A	8/9	Update to table of contents
N/A	10	Minor correction to glossary
1	12	Update to trial summary: revised recruitment figure and study duration
1	13	Continuity of study duration update.
2.4	16	Update to number of GP practices in light of closures.
3	20	Minor clarification to staff responsibilities
4.1	21	Amendment now includes a specific upper time window to enable monitoring and reporting of DNA's for the Trial. This is because the screening programme's procedures allow for an unlimited number of appointments at the same scheduled screening visit. A time window allows for a minimum of 1 or 2 rescheduling of appointments before being deemed to have failed to attend that particular screening episode, as per the NDESP (National Diabetic Eye Screening Programme) QA standards.
5.1/5.2	22	Amendment to inclusion and exclusion criteria to state exclusion of patients who have only one eye.
6.1	24	Minor wording clarification to screening section.
6.2	24	Minor wording clarifications to consent section.
6.4	25	Minor wording clarifications to enrolment section.
6.5	26	Minor wording clarifications Randomisation and Back-Up Randomisation process.
7.1.1	28	Further explanation of amendment to include upper time window and appointment scheduling to enable monitoring and reporting of DNA's for the Trial – Primary outcome
8.1	30	Further explanation of amendment to include upper time window to enable monitoring and reporting of DNA's for the Trial – Primary outcome
8.5	32	Update and clarification to Workstream F procedures
9.2	34	Minor clarifications to Randomisation process
9.4	34/35	Updates to sample size calculations and justification.
11.3.1	39/40	Clarifications to assent procedure.

18.5 Version 5 (11/January/2018)

Section	Page	Amendments
N/A	N/A	Clarifications, small amendments and updates to endnotes and referencing have been made throughout the Protocol.
NA	10	Glossary updated.
1	12	Clarifications to description of population sample and Study centre information
1	12	Study duration and Schematic of Trial design updated to reflect

		extension to project duration.
2.2	14	Update to literature on back ground on diabetic retinopathy
2.3	15	Update to the literature on screening for diabetic retinopathy
2.3.1	15	Updated to reflect changes in screening practice
2.5	17	Updated to reflect changes in screening practice. Literature updates.
2.6	18	Minor corrections and clarifications.
5.3	22	Clarification to descriptions of premature discontinuations and withdrawals
6.1 – 6.5	25-27	Minor corrections and Minor clarifications; on completion of Consent and Randomisation log, enrolment into questionnaire collection and to patient enrolment process.
6.6	28	Clarification to process for patients who are assessed with a false positive result.
7.1.1	29	Clarification on recalls for patients who DNA.
8.1	31	Updated to reflect extension to project duration.
8.6	34	Minor amendments to wording of lost to follow-up examples.
9.2	36	Clarification of sample target and numbers to be recruited.
9.4	36	Minor clarifications.
9.6	38	Clarification to analysis process.

18.6 Version 6.0 (20/November/2018)

Section	Page	Amendments
2.3.1	15	Minor grammatical correction and formatting.
4.2	20	Clarifications and further explanation of outcomes. Outcome removed as this was a workstream F outcome instead of Workstream E, therefore is not applicable for the analysis of the trial.
8.2	31	Updated to provide further clarification
9.2	35	Updated to provide further clarification
9.6	37	Update to statistical considerations.

19 REFERENCES

1. Diabetes UK: Diabetes in the UK 2010: Key statistics on diabetes. 2010.
2. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL: The Wisconsin epidemiologic study of diabetic retinopathy. II. Prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years. *Arch Ophthalmol* 1984, 102:520-526.
3. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL: The Wisconsin epidemiologic study of diabetic retinopathy. III. Prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years. *Arch Ophthalmol* 1984, 102:527-532.
4. Liew. G, Michaelides. M, Bunce. C; A comparison of the causes of blindness certifications in England and Wales in working age adults (16 to 64 years), 1999 to 2000 with 2009 to 2010. *BMJ Open* 2014, 12 February 2014
5. Global Report on Diabetes . WHO http://www.who.int/diabetes/global-report/en/ accessed 11.01.2017
6. NHS diabetic eye screening programme. Diabetic eye screening: commission and provide. https://www.gov.uk/government/collections/diabetic-eye-screening-commission-and-provide accessed 11.01.2017
7. UK National Screening Committee. The UK NSC recommendation on Diabetic Retinopathy screening in adults. January 2016.Available at http://legacy.screening.nhs.uk/diabeticretinopathy.Last accessed 4 April 2017.
8. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J,Sivaprasad S et al. Development of a cost-effectiveness model foroptimisation of the screening interval in diabetic retinopathy screening.*Health Technol Assess* 2015; 19: 1–116.
9. Taylor-Phillips S, Mistry H, Leslie R, Todkill D, Tsertsvadze A,Connock M et al. Extending the diabetic retinopathy screeninginterval beyond 1 year: systematic review. *Br J Ophthalmol* 2016;100: 105–114.
10. Leese GP, Stratton IM, Land M, Bachmann MO, Jones C, ScanlonP et al. Progression of diabetes retinal status within community screening programs and potential implications for screening intervals.*Diab Care* 2015; 38: 488–494.
11. Younis N, Broadbent DM, Vora JP, Harding SP: Incidence of sight-threatening retinopathy in patients with type 2 diabetes in the Liverpool Diabetic Eye Study: a cohort study. *Lancet* 2003, 361:195-200.
12. Younis N, Broadbent DM, Harding SP, Vora JP: Incidence of sight-threatening retinopathy in Type 1 diabetes in a systematic screening programme. *Diabet Med* 2003, 20:758-765.
13. Younis N, Broadbent DM, Harding SP, Vora JP: Prevalence of diabetic eye disease in patients entering a systematic primary care-based eye screening programme. *Diabet Med 2002,* 19:1014–1021
14. Kohner EM, Stratton IM, Aldington SJ, Holman RR, Matthews DR: Relationship between the severity of retinopathy and progression to photocoagulation in patients with Type 2 diabetes mellitus in the UKPDS (UKPDS 52). *Diabetic Medicine* 2001, 18:178-184.
15. The Diabetes Control and Complications Trial Research Group: The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993, 329:977-986.
16. Davis MD, Fisher MR, Gangnon RE, Barton F, Aiello LM, Chew EY, Ferris FL, 3rd, Knatterud GL: Risk factors for high-risk proliferative diabetic

retinopathy and severe visual loss: Early Treatment Diabetic Retinopathy Study Report #18. *Invest Ophthalmol Vis Sci* 1998, 39:233-252.

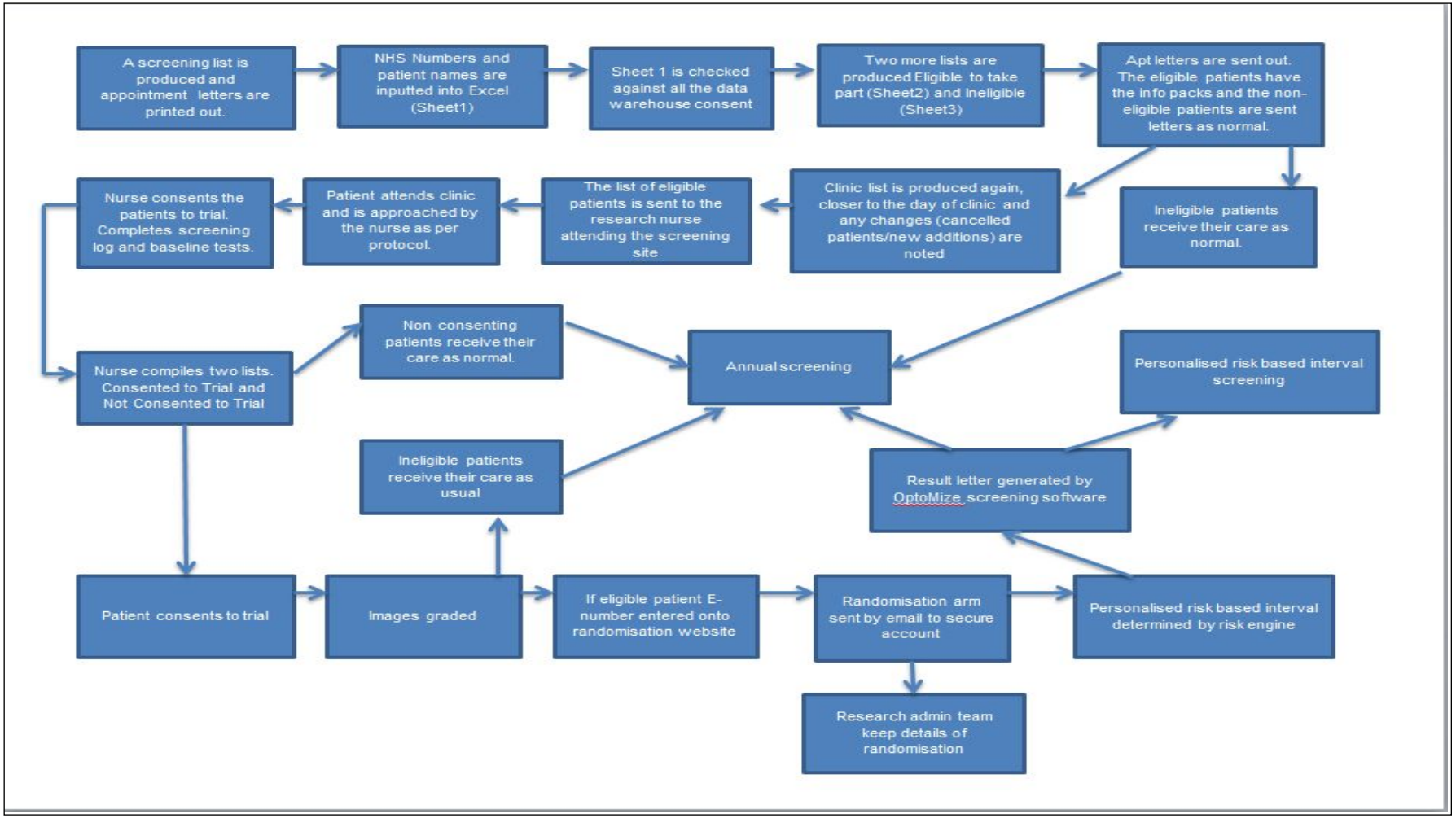
17. Marshall G, Garg SK, Jackson WE, Holmes DL, Chase HP: Factors influencing the onset and progression of diabetic retinopathy in subjects with insulin-dependent diabetes mellitus. *Ophthalmology* 1993, 100:1133-1139.
18. Stratton IM, Kohner EM, Aldington SJ, Turner RC, Holman RR, Manley SE, Matthews DR: UKPDS 50: risk factors for incidence and progression of retinopathy in Type II diabetes over 6 years from diagnosis. *Diabetologia* 2001, 44:156-163.
19. Agardh CD, Agardh E, Torffvit O: The association between retinopathy, nephropathy, cardiovascular disease and long-term metabolic control in type 1 diabetes mellitus: a 5 year follow-up study of 442 adult patients in routine care. *Diabetes Res Clin Pract* 1997, 35:113-121.
20. Vijan S, Hofer TP, Hayward RA: Estimated Benefits of Glycemic Control in Microvascular Complications in Type 2 Diabetes. *Annals of Internal Medicine* 1997, 127:788-795.
21. UK Prospective Diabetes Study Group: Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. . *BMJ* 1998, 317:703-713.
22. Gallego PH, Craig ME, Hing S, Donaghue KC: Role of blood pressure in development of early retinopathy in adolescents with type 1 diabetes: prospective cohort study. *BMJ* 2008, 337:a918.
23. Chaturvedi N, Sjolie AK, Stephenson JM, Abrahamian H, Keipes M, Castellarin A, Rogulja-Pepeonik Z, Fuller JH: Effect of lisinopril on progression of retinopathy in normotensive people with type 1 diabetes. The EUCLID Study Group. EURODIAB Controlled Trial of Lisinopril in Insulin-Dependent Diabetes Mellitus. *Lancet* 1998, 351:28-31.
24. Keech A, Simes RJ, Barter P, Best J, Scott R, Taskinen MR, Forder P, Pillai A, Davis T, Glasziou P, et al: Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. *Lancet* 2005, 366:1849-1861.
25. Klein R, Klein BEK, Moss SE, Cruickshanks KJ: The Wisconsin epidemiologic study of diabetic retinopathy: XVII: The 14-year incidence and progression of diabetic retinopathy and associated risk factors in type 1 diabetes. *Ophthalmology* 1998, 105:1801-1815.
26. Aspelund T, Porisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, Einarsson S, Palsson O, Bek T, Stefansson E: Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. *Diabetologia* 2011, DOI 10.1007/s00125-011-2257-7.
27. Javitt JC, Aiello LP: Cost-Effectiveness of Detecting and Treating Diabetic Retinopathy. *Annals of Internal Medicine* 1996, 124:164-169.
28. Jones S, Edwards RT. Diabetic retinopathy screening: a systematic review of economic evidence. Diabet Med 2010, 27(3):249-56. DOI: 10.1111/j.1464-5491.2009.02870.x
29. James M, Turner DA, Broadbent DM, Vora J, Harding SP. Cost effectiveness analysis of screening for sight threatening diabetic eye disease. BMJ 2000, 320(7250):1627-31
30. Davies R SP, Canning C,: Simulation of diabetic eye disease to compare screening policies. *British Journal of Ophthalmology* 1996, 80:945-950.
31. Stefansson E, Bek T, Porta M, Larsen N, Kristinsson JK, Agardh E. Screening and prevention of diabetic blindness. *Acta Ophthalmol Scand* 2000;78:374-385

32. Agardh E, Tabayat-Khani P. Adopting 3-year screening intervals for sight-threatening retinal vascular lesions in type 2 diabetic subjects without retinopathy. *Diab Care* 2011;34:1-2
33. Stratton IM, Aldington SJ, Taylor DJ, Adler AI, Scanlon PH. A simple risk stratification for time to development of sight-threatening diabetic retinopathy. *Diab Care*. 2013 Mar; 36(3):580-5. doi: 10.2337/dc12-0625. Epub 2012 Nov 12.
34. Scanlon PH, Stratton IM, Histed M, Chave SJ, Aldington SJ. The influence of background diabetic retinopathy in the second eye on rates of progression of diabetic retinopathy between 2005 and 2010. *Acta Ophthalmol.* 2013 Aug; 91(5):e335-9. doi: 10.1111/aos.12074. Epub 2013 Apr 1.
35. Vijan S, Hofer TP, Hayward RA: Cost-Utility Analysis of Screening Intervals for Diabetic Retinopathy in Patients With Type 2 Diabetes Mellitus. *JAMA: The Journal of the American Medical Association* 2000, 283**:**889-896.
36. Yeo ST, Edwards RT, Luzio SD, Charles JM, Thomas RL, et al. Diabetic retinopathy screening: perspectives of people with diabetes, screening intervals and costs of attending screening. *Diabet Med* 2012, 29(7):878-85. DOI: 10.1111/j.1464-5491.2012.03637.x
37. Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, 1727–36. doi:10.1007/s11136-011-9903-x
38. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics* 1995, 7(6):490-502. DOI:10.2165/00019053-199507060-00004
39. Machin D, Campbell M, Fayers P, Pinol A. Sample Size Tables for Clinical Studies. Blackwell Science, 2nd edition, Oxford. 1997

20 DOCUMENTS SUPPLEMENTARY TO THE PROTOCOL

21 APPENDICES

Appendix 1. Consent Pathway



CONSORT Statement 2006 - Checklist for Non-inferiority and Equivalence Trials

Items to include when reporting a non-inferiority or equivalence randomized trial

PAPER SECTION And topic	Item	Descriptor	Reported on Page #
TITLE & ABSTRACT	1	How participants were allocated to interventions (*e.g.*, "random allocation", "randomized", or "randomly assigned"), *specifying that the trial is a non-inferiority or equivalence trial.* Title: "Safety, efficacy and cost effectiveness of individualised screening for diabetic retinopathy: the ISDR open label, equivalence randomised controlled trial"	1,2
INTRODUCTION Background	2	Scientific background and explanation of rationale, *including the rationale for using a non-inferiority or equivalence design.*	3
METHODS Participants	3	Eligibility criteria for participants *(detailing whether participants in the non-inferiority or equivalence trial are similar to those in any trial(s) that established efficacy of the reference treatment)* and the settings and locations where the data were collected.	3
Interventions	4	Precise details of the interventions intended for each group *detailing whether the reference treatment in the non-inferiority or equivalence trial is identical (or very similar) to that in any trial(s) that established efficacy,* and how and when they were actually administered. How the interventions were administered is described in the paper. No other trial has been conducted.	4
Objectives	5	Specific objectives and hypotheses, *including the hypothesis concerning non-inferiority or equivalence.*	4
Outcomes	6	Clearly defined primary and secondary outcome measures *detailing whether the outcomes in the non-inferiority or equivalence trial are identical (or very similar) to those in any trial(s) that established efficacy of the reference treatment* and, when applicable, any methods used to enhance the quality of measurements (*e.g.*, multiple observations, training of assessors).	4
Sample size	7	How sample size was determined *detailing whether it was calculated using a non-inferiority or equivalence criterion and specifying the margin of equivalence with the rationale for its choice.* When applicable, explanation of any interim analyses and stopping rules (*and whether related to a non-inferiority or equivalence hypothesis*). Justification of the sample size is provided (see also supplementary appendix). There was no interim analysis or stopping rules in the study, however if the IDSMC or TSC had any concerns with data or safety they could recommend that the study be stopped	4
Randomization -- Sequence generation	8	Method used to generate the random allocation sequence, including details of any restrictions (*e.g.*, blocking, stratification)	3
Randomization -- Allocation concealment	9	Method used to implement the random allocation sequence (*e.g.*, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.	3
Randomization -- Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.	3
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated. This was an open label trial. Screening staff and assessors were observer-masked to intervention arm, risk calculation and interval.	3
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s), *specifying whether a one or two-sided confidence interval approach was used.* Methods for additional analyses, such as subgroup analyses and adjusted analyses.	5

RESULTS Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.	Figure 1 shows the flow of participants. Protocol deviations are explained on pages 5 and 6
Recruitment	14	Dates defining the periods of recruitment and follow-up.	5
Baseline data	15	Baseline demographic and clinical characteristics of each group.	Table 1; Supplementary Table 2; Results section (Page 6)
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was *"intention-to-treat" and/or alternative analyses were conducted*. State the results in absolute numbers when feasible (*e.g.*, 10/20, not 50%).	Figure 1 and Pages 5-6
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (*e.g.*, 95% confidence interval). *For the outcome(s) for which non-inferiority or equivalence is hypothesized, a figure showing confidence intervals and margins of equivalence may be useful.*	Pages 6-7, Figure 2 and Table 2 and ESM
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.	6
Adverse events	19	All important adverse events or side effects in each intervention group. Side effects are not applicable	6
DISCUSSION Interpretation	20	Interpretation of the results, taking into account the *non-inferiority or equivalence hypothesis and any other* study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.	7-9
Generalizability	21	Generalizability (external validity) of the trial findings.	7-8
Overall evidence	22	General interpretation of the results in the context of current evidence.	8-9

CHEERS Checklist
Items to include when reporting economic evaluations of health interventions

The **ISPOR CHEERS Task Force Report**, *Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluations Publication Guidelines Good Reporting Practices Task Force,* provides examples and further discussion of the 24-item CHEERS Checklist and the CHEERS Statement. It may be accessed via the *Value in Health* or via the ISPOR Health Economic Evaluation Publication Guidelines – CHEERS: Good Reporting Practices webpage: http://www.ispor.org/TaskForces/EconomicPubGuidelines.asp

Section/item	Item No	Recommendation	Reported on page No/ line No
Title and abstract			
Title	1	Identify the study as an economic evaluation or use more specific terms such as "cost-effectiveness analysis", and describe the interventions compared.	1
Abstract	2	Provide a structured summary of objectives, perspective, setting, methods (including study design and inputs), results (including base case and uncertainty analyses), and conclusions.	1,2
Introduction			
Background and objectives	3	Provide an explicit statement of the broader context for the study. Present the study question and its relevance for health policy or practice decisions.	3
Methods			
Target population and subgroups	4	Describe characteristics of the base case population and subgroups analysed, including why they were chosen.	3-4
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.	3-4
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.	4-5
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.	4
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	5 + ESM
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.	5
Choice of health outcomes	10	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.	5 + ESM
Measurement of effectiveness	11a	*Single study-based estimates:* Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data.	5

	11b	*Synthesis-based estimates:* Describe fully the methods used for identification of included studies and synthesis of clinical effectiveness data.	N/A
Measurement and valuation of preference based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.	5 + ESM
Estimating resources and costs	13a	*Single study-based economic evaluation:* Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	5 + ESM
	13b	*Model-based economic evaluation:* Describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	N/A
Currency, price date, and conversion	14	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.	ESM
Choice of model	15	Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended.	ESM
Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytical model.	ESM
Analytical methods	17	Describe all analytical methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (such as half cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.	ESM
Results			
Study parameters	18	Report the values, ranges, references, and, if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate. Providing a table to show the input values is strongly recommended.	Supplementary tables 11 - 13
Incremental costs and outcomes	19	For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios.	Table 3
Characterising uncertainty	20a	*Single study-based economic evaluation:* Describe the effects of sampling uncertainty for the estimated incremental cost and incremental effectiveness parameters, together with the impact	ESM

		of methodological assumptions (such as discount rate, study perspective).	
	20b	*Model-based economic evaluation:* Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions.	N/A
Characterising heterogeneity	21	If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information.	N/A
Discussion			
Study findings, limitations, generalisability, and current knowledge	22	Summarise key study findings and describe how they support the conclusions reached. Discuss limitations and the generalisability of the findings and how the findings fit with current knowledge.	7-9 + ESM
Other			
Source of funding	23	Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other non-monetary sources of support.	9
Conflicts of interest	24	Describe any potential for conflict of interest of study contributors in accordance with journal policy. In the absence of a journal policy, we recommend authors comply with International Committee of Medical Journal Editors recommendations.	9

For consistency, the CHEERS Statement checklist format is based on the format of the CONSORT statement checklist

The **ISPOR CHEERS Task Force Report** provides examples and further discussion of the 24-item CHEERS Checklist and the CHEERS Statement. It may be accessed via the *Value in Health* link or via the ISPOR Health Economic Evaluation Publication Guidelines – CHEERS: Good Reporting Practices webpage: http://www.ispor.org/TaskForces/EconomicPubGuidelines.asp

The citation for the CHEERS Task Force Report is:
Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS)—Explanation and elaboration: A report of the ISPOR health economic evaluations publication guidelines good reporting practices task force. Value Health 2013;16:231-50.

Web appendix | DELTA[2] recommended reporting items for the sample size calculation of a randomised controlled trial with a superiority question

Recommended reporting items	Page and line numbers where item is reported
Core items	
(1) Primary outcome (and any other outcome on which the calculation is based)	page 4
If a primary outcome is not used as the basis for the sample size calculation, state why	NA
(2) Statistical significance level and power	page 4
(3) Express the target difference according to outcome type	page 4
(a) Binary—state the target difference as an absolute or relative effect (or both), along with the intervention and control group proportions. If both an absolute and a relative difference are provided, clarify if either takes primacy in terms of the sample size calculation	page 4
(b) Continuous—state the target mean difference on the natural scale, common standard deviation, and standardised effect size (mean difference divided by the standard deviation)	NA
(c) Time-to-event—state the target difference as an absolute or relative difference (or both); provide the control group event proportion, planned length of follow-up, intervention and control group survival distributions, and accrual time (if assumptions regarding them are made). If both an absolute and relative difference are provided for a particular time point, clarify if either takes primacy in terms of the sample size calculation	NA
(4) Allocation ratio	page 4
If an unequal ratio is used, the reason for this should be stated	NA
(5) Sample size based on the assumptions as per above	page 4
(a) Reference the formula/sample size calculation approach, if standard binary, continuous, or survival outcome formulas are not used. For a time-to-event outcome, the number of events required should be stated	
(b) If any adjustments (eg, allowance for loss to follow-up, multiple testing) that alter the required sample size are incorporated, they should also be specified, referenced, and justified along with the final sample size	pages 4-5
(c) For alternative designs, additional input should be stated and justified. For example, for a cluster randomised controlled trial (or an individually randomised controlled trial with clustering), state the average cluster size and intracluster correlation coefficient(s). Variability in cluster size should be considered and, if necessary, the coefficient of variation should be incorporated into the sample size calculation. Justification for the values chosen should be given	page 4
(d) Provide details of any assessment of the sensitivity of the sample size to the inputs used	pages 4-5
Additional items for grant application and trial protocol	
(6) Underlying basis used for specifying the target difference (an important or realistic difference)	NA
(7) Explain the choice of target difference—specify and reference any formal method used or relevant previous research	NA
Additional item for trial results paper	
(8) Reference the trial protocol	page 3

This set of reporting items has been developed with the conventional statistical (Neyman-Pearson) approach to a sample size calculation in mind. Some of the reporting items would differ if another approach were to be used. This checklist has been taken from table 1 in *BMJ* 2018;363:k3750, as a standalone document for readers to print out or fill in electronically.

Response to Reviews by The Lancet July 2019

We would like to thank the reviewers for their timely, thoughtful and helpful responses to our manuscript submission. We have amended the manuscript to address the reviewers' concerns and to provide improved clarification. Our responses to the reviewers' comments have resulted in some slight adjustments to the descriptions of the methods and have refined the way in which the results are presented compared to in the original submission. We have added additional methodology and results to the supplementary file and reduced the number of figures and tables in the main file, moving them to the supplement. The messages, conclusions and findings regarding risk-based variable-interval screening remain unchanged. We have also made some changes to the references.

We have uploaded the various additional checklists. We think that they are not called out in the manuscript but would be happy to do so if requested.

Editorial team comments

1. Please indicate after each of the reviewers' points the text changes which have been made (if any) and the line number on the revised manuscript at which your change can be found. [Line numbers can be added to your word document using the 'page layout' tab. Please select continuous numbers. **Authors' responses in bold [Continuous line numbers added to the tracked version of the manuscript].**

2. Please indicate any authors who are full professors. **Added to manuscript**

3. Please list the highest degree for each author (one degree only, please). **Inserted**

4. Please check that all author name spellings and affiliations are correct. **Checked**

5. For randomised trials please follow the CONSORT reporting guidelines and CONSORT for abstracts and include a CONSORT checklist with your resubmission**. Updated, including in response to reviewer 1.**

6. Please ensure that the title of the paper is non-declamatory (i.e. it describes the aim of study rather than the findings) and that it includes a description of the study type (e.g. a randomised controlled trial). **Revised as per reviewer 1**

7. Please limit the summary to pre-defined primary endpoints and safety endpoints. **Done**

8. For RCTs, please state the trial registration number **ISRCTN 87561257 already included at line 48.**

9. At the end of the methods section please state the role of the funder in: data collection, analysis, interpretation, writing of the manuscript and the decision to submit. Please also state which author(s) had access to all the data, and which author(s) were responsible for the decision to submit the manuscript etc. [**Lines 232-235 and 445-447].**

10. Please explain any deviations from the protocol. **Described in the results section [lines 343-345].**

11. Please report all outcomes specified in the protocol. **We now refer to the list of secondary outcomes as specified in the protocol and have added a supplementary table 9 (see full response to Q17-18).**

12. If any exploratory outcomes are reported that were not pre-specified, please make it clear that these analyses were post-hoc.

We conducted three post-hoc analyses, one included in the original submission on HbA1c to assess the potential effect on glycaemic control of risk based screening (clarification added at line 295) and two in response to reviewer comments, attendance rates over 24 months [271-273] and switch in intervals in the individualised arm [273-276] (comments 36 and 43 respectively).

13. Please use rINNs for drug names. For genes and proteins, authors can use their preferred terminology so long as it is in current use by the community, but should provide the preferred human name from Uniprot (http://www.uniprot.org/uniprot/) for proteins and HUGO (http://www.genenames.org) for genes at first use to assist non-specialists – **N/A**

14. For drug studies, please ensure that details of doses, route of delivery, and schedule are included. **N/A**

15. For the main outcome measures, please include a result for each group, plus a point estimate (e.g. RR, HR) with a measure of precision (e.g. 95% CI) for the absolute difference between groups, in both the Summary and the main Results section of the paper. **Done**

16. p-values should be exact, but no longer than 4 decimal places (e.g. p<0.0001). Two decimals are acceptable in tables for non-significant p-values**. Done**

17. Please provide absolute numbers to accompany all percentages. Percentages should be rounded to whole numbers unless the study population is very large (>10 000 individuals). **Done**

18. Please give 95% confidence intervals for hazard ratios/odds ratios. **N/A**

19. For means, please provide standard deviation (or error, as appropriate). **Done**

20. Please provide interquartile ranges for medians. **Done**

21. Please provide numbers at risk for Kaplan-Meier plots and ensure that plots include a measure of effect (eg, log-rank p); estimates should be reported with 95% CIs. **N/A**

22. Please ensure that the Discussion contains a section on limitations of the study. **Done**

23. Please provide the text, tables, and figures in an editable format. See link above this list for details of acceptable formats for figure files. **Done**

24. Our production system is not compatible with Endnotes. Please convert to normal text. **N/A**

25. If accepted, only 5-6 non-text items (figures, tables, or panels) can be accommodated in the print edition; additional material can be provided in a web appendix. Please indicate which items can go in a web appendix.

Reduced to 6 items. We have combined our Tables 2 and 3 into a new Table 2 (as suggested by Reviewer #1), moved the previous Table 4 into the supplementary file now at Supplementary Tables 11 and 12 and revised the previous Table 5, now Table 3.

26. Please provide a research in context panel with 3 parts: Evidence before this study (which includes a description of how you searched for evidence and how you assessed the quality of that evidence); Added value of the study; and Implications of all the available evidence. **Done**

27. At the end of the manuscript, please summarise the contribution of each author to the work. **Done**

28. At the end of the manuscript please summarise the declaration of interests for each author**. Done**

29. If you have not yet done so, please return all signed authorship statements and conflict of interest forms. We also require signed statements from any named person in the acknowledgements saying that they agree to be acknowledged. **Re-submitted with new title**

30. For any personal communication, please provide a letter showing that the person agrees to their name being used. **N/A**

31. As corresponding author, please confirm that all authors have seen and approved of the final text. **Done**

32. If your author line includes a study group, collaborators' names and affiliations may be listed at the end of the paper or in the appendix. Additionally, if you wish the names of collaborators within a study group to appear on PubMed, please upload with your revision a list of names of all study group members presented as a two-column table in Word. First and middle names or initials should be placed in the first column, and surnames in the second column. Names should be ordered as you wish them to appear on PubMed. The table will not be included in the paper itself - it's simply used to make sure that PubMed adds the names correctly. **Done. Uploaded with re-submission**

33. Please note our guideline length for research articles is 3500 words and 30 references. For RCTs, the text can be expanded to 4500 words. **Word count, refs Add statement justifying increased length.**

34. From July 1, 2018, all submitted reports of clinical trials must contain a data sharing statement, to be included at the end of the manuscript or in an appendix (please provide as a separate pdf). Data sharing statements must indicate:
*Whether data collected for the study, including individual participant data and a data dictionary defining each field in the set, will be made available to others;
*What data will be made available (identified participant data, participant data with identifiers, data dictionary, or other specified data set);
*Whether additional, related documents will be available (e.g. study protocol, statistical analysis plan, informed consent form);
*When these data will be available (beginning and end date, or "with publication", as applicable);
*Where the data will be made available (including complete URLs or email addresses if relevant);
*By what access criteria data will be shared (including with whom, for what types of analyses, by what mechanism – e.g, with or without investigator support, after approval of a proposal, with a signed data access agreement - or any additional restrictions).

Reviewed and additional detail added at lines 449-453.

Reviewers' comments

Reviewer #1

THELANCET-D-19-02765_response Broadbent et al.

In the main the study is well conducted, reported and presented. It is a complicated study and the comments do add up but

Major comments

1. When doing an equivalence trial there still is a wish to show some benefit for the new intervention. Page 18 of the protocol summarises things nicely. The introduction, research in context and discussion should be in terms of the benefits and risks for the intervention

Thank you these thoughtful comment. We agree that our messages have focussed mainly on safety and harms. In the protocol we identified, including in discussion with our PPI group, 4 general areas of potential benefit:

- **"Reduced burden of appointments for people who are at low risk of developing STDR." We have discussed at several locations in the results and discussion**
- **"Reduced anxiety for people with diabetes with reassurance that personalised risk is being determined." We have shown the study is safe and have addressed this in the conclusion [lines 420-422].**
- **"Earlier detection of STDR in people at high risk associated with better outcomes in terms of vision and need for fewer and less invasive treatments." Here we have identified the earlier detection of STDR in people of high risk [lines 305-307, 376-378, 409-412] but didn't find a reduction in invasive treatments, at least within 6 months of STDR being detected [lines 281-282]. We were not able to identify better outcomes in terms of vision but didn't detect any difference in VA or VI [lines 291-293].**
- **"Improved cost-effectiveness through redirection of resource of the screening programme, including increased capacity to see newly diagnosed patients." We have covered this in several places throughout the manuscript.**

We have added a sentence at lines 410-413 to highlight this important feature of an equivalence study and emphasise the other potential benefits over and above safety, "A number of benefits include a lower burden of appointments, earlier detection of STDR for high risk people and increased capacity to see people when they are newly diagnosed."

2. Can a CONSORT checklist for non-inferiority and equivalence please be completed? **Done**

a. In the main the checklist is similar but there needs to be a justification as to why the study was set up as an equivalence trial. There are issues as to choice of control as well but this is self-explanatory.

Thank you for this comment. A consort checklist for non-inferiority and equivalence trials has been completed. We have added a sentence to the introduction, [lines 112-114], "An equivalence design was selected, as opposed to a superiority trial, because the aim was to demonstrate equivalence between attendance rates rather than a difference." One of the worries about extending screening intervals was that those given an interval >1 year wouldn't return. There was nothing to suggest that they would be <u>more</u> likely to attend if given a longer interval.

b. Can the study title please give the study design: single centre, open label, equivalence trial?

Changed to "Safety, efficacy and cost effectiveness of individualised screening for diabetic retinopathy: the ISDR single centre, open label equivalence randomised controlled trial."

c. Following on from the study is open label study and should be described as such including in the study title **See above**

d. Masked clinical assessors I would suggest better phrased as observer-blind.

We don't typically use "blind" in ophthalmology studies so changed to "observer-masked" [Line 142].

3. Under weaknesses can it please be given that this is a single centre study. **Inserted, "Ours is a single programme …" to line 338.**

How generalizable is this study to other centres? Particularly centres which are not 95% white and are greater than 2% Asian?

We recognise that this is a weakness and have discussed in lines 338-342. We have added the phrase "… but our results should be treated with caution in areas with higher prevalence, poorer control of diabetes, wider ethnic group representation, or in programmes in set-up."

4. Can you please complete the DELTA2 sample size checklist recently published and given in Table 2 in (https://doi.org/10.1136/bmj.k3750). The focus in this is on superiority trials but the same principles apply. **Completed and uploaded.**

a. I cannot replicate the sample size calculation as neither in the paper or protocol the baseline risk given and so one does not know what 5% is an increase of (the detail is in the protocol)

We thank this reviewer for raising this question. We have now clarified this aspect in the paper. The corresponding sentence in the paper [line 191-192] now reads: "A conservative expected attendance rate of just below 75% (74.6%) was considered from a range of previously observed local values".

We have added further detail on the sample size calculation in a new section, "Sample size calculation" in the supplementary file. As indicated in the protocol, due to a lower recruitment rate and a lower attendance rate than expected, the sample size required was revisited in November 2015. We considered a range of possible values for the recruitment rates based on the recent figures of attendance observed at the time (just below 75%). We were aware that very small deviations from the initial 75% expected recruitment rate, for example ±0.4%, would result in differences in sample size of about ± 90 patients, and consequently decided to adopt a conservative value for the expected rate to cover a range of possible scenarios. We used Equation (9.12) [1] with allocation ratio φ=1, 10% two-sided type II error ($z_{0.95}$), 2.5% one-sided type I error ($z_{0.975}$), equivalence margin=0.05 and equal attendance rates $\pi_1=\pi_1=\pi$=0.746. The sample size obtained divided by (0.94)^2 to allow for 6% annual loss in 24 months resulted in n= 4460.

[1] Machin D, Campbell M, Fayers P, Pinol A. Sample Size Tables for Clinical Studies. Blackwell Science, 3rd edition, Oxford. 1997

b. The observed rates are greater than anticipated for the primary endpoint it seems which needs to be discussed

In the protocol we pointed out that data from the Liverpool retinal screening programme suggested an attendance rate for annual follow-up screening of around 85%. However, nationally, estimates of attendance rates were closer to 75%. In the manuscript, and in the supplementary file, we now clarify

that we covered a range of possible values for the attendance rate, and that a conservative expected attendance rate of just below 75% was adopted so that a power of 90% could be achieved [line 189-192].

c. Why is 5% being used as an equivalence limit. What is the justification for this?

The equivalence margin 5% was regarded by the clinical investigators as a minimum difference in attendance rates acceptable for implementation of the risk-based variable-interval approach into clinical practice and based on experience in other equivalence and inferiority trials that have resulted in changes in clinical practice. This justification has now been added in the supplementary file in a new section, "Sample size calculation".

d. Being an equivalence trial the study has a one tailed Type I error of 2.5% (not 5% as stated). **This is now clarified at line 189. Thank you.**

e. What is the justification of the 1.5% non-inferiority limit?

We discussed the non-inferiority limit margin extensively during the trial design. We recognised that because of the relatively low incidence of STDR there was going to have to be a pragmatic trade-off between the size of the margin and the statistical power. The clinician investigators felt that a 1.5% margin on an expected 6% detection rate was reasonable for this secondary outcome.

f. As with the comment above what are the baseline risks that are being assumed

We would like to highlight that the study was powered to address the primary question (equivalence for attendance), and that the proposed sample size was put into context of the secondary outcome (STDR) for completeness. The STDR detection rate predicted for the usual care pathway during the two-year follow-up was considered to be approximately 6%, based on previous figures from the Liverpool (National) Diabetic Eye Screening Programme (LDESP).

g. From my read of the paper there does not seem to be any hierarchical procedure to go from equivalence to non-inferiority

We are not certain about the reviewer's question here. We wonder if we have mistakenly suggested that we viewed the secondary safety analyses as a progression from the primary analysis. The equivalence and non-inferiority relate to different analyses. We would be happy to provide a further response to this question if the reviewer wishes.

h. Can the fact the study had a sample size re-estimation

We are unsure whether this is a question or a remark. This reviewer is correct in that there was a sample size re-estimation early in the RCT as described in the protocol (section 9.4), although this did not change the study design. The motivation for recalculation was the slower than expected recruitment rate.

5. For the statistical analysis

a. Can greater detail please be provided as to the statement about simple imputation (can be supplemental material)

Following this suggestion we have now added the statistical analyses involving simple imputation using the most and least conservative scenarios as a new Supplementary Table 4. They were described in the protocol and included in the SAP and full statistical report.

b. Following on from the statement above can a statement please be added that there is no allowance for multiplicity?

We have added at line 208 the following: "There was no allowance for multiplicity".

c. For HE the bootstrapping how many bootstrap samples were done. Was an assessment of the stability of the estimates undertaken be replicating the analysis 5-6 times (and upping the replications if not stable)?

In the new health economics section in the supplementary file, subsection "Cost effectiveness" we have added "estimates following 1000 iteration bootstraps were unstable; increasing this number to 10,000 iterations correctly produced estimates which converged to the means of true sampling population".

d. It is not clear within the paper how CI are calculated for the binary outcome. For the primary outcome the simple Normal approximation may be fine. For the non-inferiority assessment Normal approximation should not be as the Normal approximation does not hold due to the low response rates. I would suggest Wilson Score CI for the primary outcome (could apply to both outcomes)

We agree that in situations where the response rate is expected to be low, the normal approximation may not work well. This is especially the case for relatively low sample sizes. Given that we are here dealing with large sample sizes, we followed the Central Limit Theorem, which states that the sampling distribution of the sample means approaches a normal distribution as the sample size gets larger, and this is regardless of the shape of the population distribution. Nevertheless, we acknowledge that this is an assumption that should have been checked, and following this reviewer's remark we now report CIs based on the Wilson score (Newcombe score) for our non-inferiority analysis. As expected, the CIs were very similar, and the interpretation did not change. The final ISDR statistical report will capture these calculations as post-hoc analyses. We also checked the primary outcome and confirmed the CIs are the same regardless of whether the calculation is based on the normal distribution or the Wilson score.

6. Can the statistical analysis plan for this study can it please be provided? **A web link was provided in the manuscript as reference 17. We have uploaded the SAP with the revised submission.**

7. I am not a health economist but I do not fully understand the health economic results.

a. The protocol says that ICERs will be produced and CEACs but no CEACs are provided and no ICER as what I would understand given.

Thank you for these comments. The original intent had been to produce CEAC and ICERs. We now present ICERs and 95% Cis in a more interpretable format in table 3. Since some ICERs were negative we feel that the alternative statistic of incremental net monetary benefit is a more useful figure to report for decision makers. Since our treatment is dominant in cost saving and the majority of bootstrap iterations showed a positive incremental QALY, CEACs were uninformative as they showed our treatment is almost always cost effective. We undertook considerable multiple imputation modelling and the data shows that individualised risk-based variable-interval screening was strictly dominant across all relevant thresholds, however we believe that

the CEAC would have been misleading to present given the small size of differences in quality of life, as such we focus upon possible cost savings given QALYs were not diminished between arms. This is now discussed in our results section [lines 316-320] and further within the supplement.

b. The cost effectiveness plane in Figure 3 seems to be showing the intervention is less expensive and has better QALYs?

We thank you for this guidance and agree, yet our confidence intervals of the mean do cross the zero incremental threshold (despite a positive skew). We feel that the conservative message of equivalence in patient safety and outcome, but reduction in cost, are claims we are happier to make. We have taken the opportunity to add the 95% confidence boundaries to the figure.

c. There is a statement as to the strict dominance of the intervention which in context of Figure 3 seems consistent and why there is no ICER?

The ICER is now given in our revised HE table (now Table 3) alongside data on incremental net marginal benefit.

d. Personally I would not rely on P-values (on a relatively small sample size) to infer no difference on QALYs. There does seem to be QALY gain but P>0.05

We now clarify in the abstract that "Incremental QALYs per person were not statistically significant" instead of non-significant. Within our analysis we do not discuss P-values as the health economics field leans towards the irrelevance of inference, where our decisions cannot be deferred in a healthcare setting due to opportunity cost. Therefore our conclusions are those regarding highest expected mean benefit. We do observe a small QALY gain which is however not significant since it crosses the threshold to be interpreted much in the same way as relative risk ratios.

e. The footer to Table 5 refer to NICE willingness to pay thresholds but there is no mention in the text to these. ICERs, in context with NICE threshold and probability of being cost effective I would expect to see. Is the probability of being cost effective 100%? If not why not say

We completely agree with this point, and as such we discuss in the results section issues with interpretation of ICERs. We have now expanded this in a new health economics methods section in our supplementary file. We have ensured the £20,000 threshold in the new Table 3 is also iterated in the health economics methods section [line 225-227] and fully elaborated in the health economics results section of the supplementary file.

f. For CEACs are not possible then it should be stated.

We now discuss the uninformative nature of CEACs applied to our findings in the results in lines 316-318.

g. Given this is a single centre study. How would the cost effectiveness be impacted if there was a different demographic distribution?

This is a good question and it is something we wish to explore in future work.

h. What is the cost effectiveness in the high, medium and low risk sub groups?

Whilst again we appreciate that this is an interesting question, the study was not powered to undertake this subgroup analysis. The comparison was risk-based versus usual care not the cost effectiveness of screening at differing risk thresholds. As such the number in the high-risk group, we feel, is too small for us to make inference of the cost effectiveness of screening for this group or the medium risk groups in isolation. Similarly, because we do not have a direct comparison across each of these risk categories against usual care.

8. Can a CHEERS checklist please be provided?

Uploaded with re-submission

9. Can the HEAP please be provided?

A web link was provided in the manuscript as reference 17. We have uploaded the HEAP with revised submission.

10. For the DMC

a. Can the charter for DMC please be provided?

We have uploaded the Independent Data and Safety Management Committee (IDSMC) charter with resubmission. Information on the role of the IDSMC, analysis, amendments and operational bias has been added to the supplementary file in a new Supplementary Methods Section.

b. Did the DMC just assess safety only i.e. did they have an efficacy or futility stopping rules.

Text included in new Supplementary Methods Section, "There were no stopping rules in the study, however if the IDSMC or TSC had any concerns with data or safety they could recommend that the study be stopped."

c. Who undertook the analysis for the DMC? **The trial statistician undertook the analysis for the IDSMC. We have included text included in new Supplementary Methods Section.** How was operational bias minimised? **We have included new text included in new Supplementary Methods Section.**

d. Were there any amendments at the request of the DMC? I could not see any in the list of amendments.

No amendments other than further clarification on information and additional details in the report. We have included text in new Supplementary Methods Section.

e. Can the frequency the committee met please be given in the paper.

The Committee met on an annual basis as agreed at the first meeting and in line with recruitment and follow-up data; text added at Line 173 and included in new Supplementary Methods Section.

11. For Table 1 and Table 2

a. Can the analysis population please be provided in the table title? **Added**

b. I would suggest that the table be on the per protocol population - with ITT in supplemental material

Thank you. The old Table 1 (ITT dataset) has been moved to the supplementary file (Supplementary Table 2) and the new Table 1 based on the PP dataset. Given that Figure 2 shows the results for both ITT and PP analyses, we have kept the ITT results in Table 2 for completeness, and have merged Tables 2 and 3 into one table following the helpful suggestion below (Comment 12).

12. Could Table 2 and Table 3 be merged? **Tables 2 and 3 are now merged into Table 2**

13. For Table 4

a. Can the analysis population please be provided in the table title**? This has been amended.**

b. Can there please be more consistency in the decimal places. **Corrected for all main and supplementary tables.**

c. Does the multiple imputation columns need to be in this table? Can these be in supplemental? **Multiple imputation statistics have been moved to the supplement.**

14. For Table 5

a. Can the analysis population please be provided in the title? **Revised.**

b. Does the multiple imputation results be in supplemental?

These have been amended and moved to the supplementary file. Our new table 3 (and supplementary tables) now specifies our intention-to-treat population and their respective values of n.

c. I would restructure the table to be more like Table 2 in format with treatments in columns and the point estimates and CI for the treatment difference.

Thank you for your advice with which we agree with. We have restructured the health economic data into a single table (new Table 3) within the paper and moved other statistics into a new health economics section in the supplementary file.

d. The CIs are quoted in text and in the abstract and so need to be in this table.

The new table now states the 95% confidence intervals. We thank the reviewer for this helpful feedback and feel this new table reflects more clearly the analysis undertaken.

15. Are Table 4 and table 5 required in the actual paper? Can one be moved to supplemental material? **See responses above.**

16. Can the CONSORT checklist for Harms please be completed? **Done**

17. Can there be an additional table giving the secondary efficacy outcomes consistent with p20 and p31. **See below.**

18. Can there be a table of the safety outcomes consistent with p31. Some of the outcomes are assessed as efficacy are safety but there are additional safety outcomes not given which should be given that safety is important for this study (and there was a DSMB). Even if there was zero events (heart attacks and strokes) then pleases still provide. Please give point estimates and confidence intervals.

THELANCET-D-19-02765_response Broadbent et al.

We respond to these important comments 17 and 18 together. Due to the nature of ISDR, some outcomes refer to both safety and efficacy, and the line between the two is not easy to draw. For example detection of STDR provides evidence of safety through disease missed and efficacy through numbers of screen episodes required to detect STDR. STDR was a key patient related outcome identified by the PPI group.

We have added a table of secondary outcomes listed in page 20 as a new Supplementary Table 8 in the supplementary file. On page 31 of the protocol we listed several variables to be collected during the trial to allow calculation of the secondary outcomes and some additional efficacy and safety outcomes.

Two of the secondary safety outcome variables have been evaluated in our analysis in methods slightly differently to the description in the protocol: we have reported treatment rates in both arms in the 3 and 6 months after diagnosis of STDR; we have reported mean numbers of visual impairment (VI) rather than new VI.

The following secondary safety outcomes proved very difficult to collect and should have been removed during one of the protocol amendments. This was an oversight and will be addressed in the report to the NIHR:

- **number of dedicated diabetes assessment clinic appointments (biomicroscopy)**
- **number of other eye appointments for DR**
- **hospital attendance for diabetic life-threatening events (heart attack and stroke) in both arms of the trial.**
- **attendance at GP annual review rates in both arms of the trial.**

We have captured the above explanations in the expanded supplementary file.

19. How easily implementable would be the risk calculation tool?

The risk calculation engine (RCE) tool is fully automated. For implementation in the research setting it required a clinical data warehouse linked to primary and secondary care databases. A bolt on piece of software developed with EMIS Health (OptoMize) transfers data to and from the local screening programme and the RCE. Currently few screening programmes have access to any information other than the screening data so implementation would need clinical information available to the screening programmes. However we have shown that by engaging GPs, patients and the Clinical Commissioning Group this has been possible. We have described the details of the risk calculation engine in Eleuteri et al, Diabetologia 2017 Ref 18. For further implementation some further development would be required. We have addressed this in the last paragraph of the research in context, "Scale up outside a research setting with further monitoring is recommended" [line 89]. We have added "Some systems development from our research setting to an implementation environment will be required "at lines 371-372 in the discussion.

Other comments

20. The protocol describes the intervention arm well: risk based screening intervals which I would suggest use in the paper.

THELANCET-D-19-02765_response Broadbent et al.

We spent a lot of time debating the correct term. We have used the term "individualised risk-based variable-interval" throughout the paper. We used "individualised" rather than "personalised" - at the time we were developing the programme of work the term "personalised medicine" was not so widely used. We believe that the terms "risk-based" and "variable-interval" are important components of the intervention.

21. Through the paper can you please avoid using the term "non-significant". Please refer to as "non-statistically significant". **We have reviewed the manuscript and corrected one iteration of this in the summary and one in the discussion**

22. As an aside the study is interesting example of when an equivalence trial is appropriate over a non-inferiority study which are in the main used now (obviously not for action). **Thank you. We agree with this.**

23. Can the data management handling and data management procedures please be described within the paper (how undertaken, systems used)? **We have added a description of the data management and handling to a new Supplementary Methods section in our supplementary file.**

24. Can more detail please be provided on the randomisation procedures? What package used, who obtained the randomisation at the study site.

We have changed this paragraph to read "Block randomisation was conducted by bespoke automated fully validated electronic allocation system in the CTRC with stratification by clinic and age using random blocks of 4 and 6 for participants aged ≥ 16 blocks and for those aged <16 years blocks of 2 to account for small numbers. The allocation sequence was generated by an independent CTRC statistician. A purpose built, dynamic data warehouse (DW) holding and linking primary and secondary care demographic, retinopathy and systemic risk factor data from 2009 populated the fields in the baseline and follow-up CRFs (OpenClinica, LLC), including data required for randomisation. Allocation was received electronically into the LDESP management system (OptoMize, EMIS Health) and participant letters generated. Screening staff and clinical assessors were observer-masked to intervention arm, risk calculation and interval" [lines 134-143].

Reviewer #3: Thank you for submitting your paper to this journal.

ISDR is clearly an important trial that potentially has far-reaching consequences in most countries where diabetic eye screening is carried out, or where it is to be implemented. The size of the RCT is truly impressive and the results are convincing for those who agreed.

25. This reviewer's main concern is the number of patients who did not agree to the trial and how this percentage has been incorporated into the economical analysis. If a substantial minority of patients cannot be convinced to give up the annual screening, it must have an effect on planning and costing of services. Or did you make the assumption that once the variable interval is within the national guidelines, then there will be no choice presented to the patients?

The plan in England is to move low risk people to 2 year screening based solely on their retinopathy levels (Scanlon et al). It has been emphasised to people with diabetes over many years that they should attend annually and so it is not surprising that without evidence on safety they were reluctant to change. It is for this reason that we undertook the RCT. Given the reassuring results of our trial it will be easier to reassure all stakeholders and people with diabetes that extended intervals are safe and more cost effective.

Reviewer #4: Broadbent and colleagues present a clinical trial looking at the consequences of moving to variable-interval screening for diabetic retinopathy in the UK. The study is an important and novel one, because diabetes is increasing globally, putting pressure on health systems to provide evidence-based care but also contain cost growth. The authors find that variable-interval screening is an attractive option, with similar rates of detection (vs. standard annual screening) and lower costs over a 2-year time horizon.

I am specifically reviewing the economic evaluation component of the study and defer to other reviewers who are experts in clinical trials to provide feedback on the rest of the study. In general, I have no major concerns about the substance of the authors' conclusions from the economic analysis, as they seem intuitive in light of the main trial's findings. However, I have a number of specific comments on how the economics section could be improved.

We thank the reviewer for these comments which were very helpful. We have taken them on board and used them to revise the manuscript text accordingly. We have added a health economics section to our supplementary file including, methods, results, new tables and a new figure. We indicate below the specific revisions made in the text.

Major comments

26. Methods section: I realize this is primarily a clinical trial article, but I do think that more information is needed on the conduct of the cost-utility analysis. The BMJ Open protocol does not contain much more information on the economics than the trial manuscript itself. The CHEERS checklist is now being used to ensure standardization of CUA/CEA reporting, and I would recommend that the authors refer to this to ensure that all the elements are mentioned. I do not think that the article needs to be longer but rather that a small text appendix plus completed CHEERS checklist would be helpful. The specific topics for which I was wanting more detail were the costing methods and the methods for constructing QALYs (where do the "appropriate utility tariffs" come from?).

We agree and have addressed this in the new health economics section to our supplementary file. We have also completed a CHEERS checklist and uploaded with the re-submission.

27. Results section: There were only 4 treatment events in the whole sample during the study period. I assume treatment to be photocoagulation (though it is not specified as far as I can tell). Two comments here. First, it does not appear that treatment costs were included in the costs reported (e.g. tables 4-5, appendix table 6). Second, I wonder whether a longer time horizon of follow up would reveal different rates of treatment in the two groups despite similar rates of positive screening. If so, then the short term cost savings reported here might not represent the difference in costs between the two patient groups over their lifecourse. I am not suggesting the study is invalid, just that it is limited in terms of what it can claim about resource savings (and I think it would be helpful for the authors to acknowledge this).

Treatment was defined as either photocoagulation, intravitreal injection of anti VEGF or vitrectomy.

What we have done relates to a further safety measure in which we looked to see if any patients required treatment within a reasonably short time period of 6 months after they screened positive. No patients required treatment within 3 months of being screen positive. Two patients developed diabetic maculopathy and required treatment within 6 months of being screen positive [lines 281-282].

We thank the reviewer for the second comment which is of course completely correct, and have now included this as a limitation in the discussion [lines 404-408]. The treatment costs are not included in this paper. We have confined the costs only to screening which were derived following an extensive

micro-costing process. We felt a fair and true representation of the spread of treatment cost could not be determined in the two year period. We hope we now make this clear in the text and in our description of the limitations of the time horizon. A health economics model is being developed which will use the trial data, incorporate a lifetime time horizon and include treatment cost estimates.

28. Results section: On a related note, the estimated QALYs in the two study arms were similar. This is not surprising, because patients were randomly allocated between the two arms, and no outcomes were measured (related to the intervention) that would be expected to significantly affect health status. Similar rates of screen positivity and treatment (low event rate noted) suggest that QALYs post-intervention will be unchanged from their baseline (presumed equivalent due to randomization) values. Without differences in QALYs, the study essentially becomes a cost-minimization study. Hence the language around "cost-effectiveness" throughout the results/discussion would probably be more accurate if it were framed as cost reduction rather than increased value for increased money (i.e, the northeast quadrant of the CE plane).

Again we thank the reviewer for this comment. A wider description of the effects as they relate to QALYs is now given in lines 311-320. We agree the reduction in cost is the important finding and have now addressed this in lines 314-321 and have taken this on board within the results. The discussion highlights the importance of the cost savings [lines 395-403].

29. Discussion: The fact that the program has been running for 3 decades and has low rates of baseline DR and progression to DR is critical from a cost-effectiveness standpoint. The cost of setting up a screening program and/or working in a higher prevalence/risk setting would be quite different. Presumably in those settings the RCE (if calibrated correctly) would still perform well at identifying case screening intervals, so would less frequent screening in the low-risk population lead to any harms? The issue here seems to be with the validity of the RCE in different settings rather than the extrapolation of this trial's main substance to other environments.

We completely agree with the reviewer; these are important points. One of the novelties and advantages of the Liverpool RCE is that it was developed using local data and as such is embedded in our local population using historical data. We have added clarification ("historical retinopathy data") to line 153. The algorithm was stable in the research setting. When used in another setting it would need to be run again using data from that setting in order to reflect the risk in the new population. As such we have tried to restrict our comments to "Evidence to date is from established programmes with relatively low rates of retinopathy." [Line 86] "…apply to other established screening programmes but should not be applied in territories without established systematic screening …" [lines 88, 415-417], and "… in established screening programmes" [line 422].

30. Discussion: Extending screening beyond 2 years is an important option to explore particularly in low- and middle-income countries and in countries with substantial rural populations - both of which have greater cost and logistical constraints than Liverpool. This would be a valuable direction for future research and replication in different settings.

This is an important point that we have tried to reflect in the discussion. A move to longer intervals in resource poor settings in low and middle income countries is attractive. However we believe that this could be unsafe where programmes are being set up or have unstable call-recall systems, where attendance is limited by factors such as travel costs and distance, or where there are high levels of prevalence. We agree that further research in these settings will be needed and until that is available strongly feel that an annual recall is the wisest option.

We have added a sentence to emphasise this, "For resource poor or rural settings in low and middle income countries further research is required before longer intervals can be contemplated." lines 358-360.

31. Table 4 has a couple of odd things. First, I notice some negative values for QALYs are being reported. Are the authors sure that QALYs are being calculated correctly? Again I can't tell because of the lack of detail in the paper/appendix. Second, on the cost section, the "societal" costs are less than the "total screening" costs. Societal costs are the sum of provider, patient, and all other costs borne, so they should be higher for either control or intervention group than the provider costs. Appendix table 6 suggests the authors are using the term societal to refer only to informal healthcare sector + non-healthcare sector costs, which is technically incorrect (see fig 1 in Sanders et al., JAMA 2016; 316:1093**).**

A small number of participants' elicited utilities were negative at each time point throughout the trial and as such produced negative QALYs, this has been discussed further within the results section of the health economics supplement. We now state the tariffs used for the EQ-5D-5L and the HUI3 [lines 221-222]. Table 4 has now been split and moved to the supplement. We apologise for the confusion surrounding our use of the term societal cost and have now corrected this to read "a societal perspective" throughout. We have defined the use of the language "social costs" in the methods section to explain these social costs comprised time lost from work of patients and companions in attending a screening episode (productivity losses) and out of pocket expenses in terms of travel and parking costs.

32. Table 5 also has a couple of odd things. First, mean societal costs (sic) as reported in the difference in means section for complete cases is positive, whereas it's the same numbers but negative for the imputation section. Is one of these a typo? Second, what is causing the ICER for complete cases for EQ5D to be GBP 14,350? Obviously this is just a point estimate, and the SE is so wide that we can't for sure say it's not cost-saving. But this does not align with the level of precision implied in the utility measures and cost measures themselves.

We apologise for having missed that positive value, it indeed should have been negative. We have now reworked, reformatted and improved the clarity of this table, which is now the only health economic summary table reported in the paper itself. When our observed incremental QALYs tended to zero our ICERs tended to infinity which skewed our means. As such we now correctly report the ICER of the mean rather than the mean ICER.

Reviewer #5: This manuscripts reports on a highly interesting randomized controlled trial, investigating the safety, efficacy and cost-effectiveness of risk-based retinopathy screening in individuals with diabetes mellitus.

The trial was designed as an equivalence study with primary endpoint attendance at first follow-up visit and various secondary endpoints. It was performed in a single centre in Liverpool.

The current manuscript reports on this primary endpoint, as well as on cases of sight threatening retinopathy detected (efficacy), cost savings, effects on quality of life and cost-effectiveness.

Thank you for these supportive comments.

Major comments

33. The cost-effectiveness study is completely piggy backed to the RCT and as such lacks a careful scoping, to justify chosen time horizon, outcome measures and costing methods. Specifically the time horizon seems inadequate in relation to the outcome measure (QALYs) chosen. For costs per QALY, a lifetime horizon would have been required.

These reviewers advise reporting on incremental costs in relation to missed cases or differences in DR scores, as being more appropriate for the time horizon chosen.

We decided to include the cost effectiveness component of our work in this manuscript because we feel that it is a key part of the interpretation of the results on safety and where one of the key messages lies. In our original submission we restricted the scoping and detail to fit with the maximum allowed word count. We have now slightly expanded the methods section of the HE methods in the main manuscript and added an extensive health economics methods section to our supplementary file.

We appreciate the comments about the short time horizon. We are working on a long-term risk based model of the cost and effects of screening and now recognise the time horizon as a limitation in the discussion of the work here reported (see responses to Reviewer 4 comment 27)

34. It is questionable to use direct measurement of quality of life for this study. Effects on quality of life will only occur once diabetic retinopathy or other complications develop. Attending screening visits has a negligible effect on quality of life as measured by HUI and EQ-5D-5L, and differences between groups with at most 18 months difference in screening interval times will be improbably to exist and impossible to detect. So here the reviewers question the correctness of the study design.

QALYs were a secondary outcome in the study but were an important component of following a cost effectiveness analysis design. The results show that the QALYs between the groups showed a marginally better outcome in the intervention arm but this finding was not statistically significant [line 310]. We do however demonstrate in the paper [lines 309-311] and in the health economics section of the supplementary file that a societal perspective yields greater cost savings for the individualised risk-based approach once the attendees and their companions' time lost from work and out of pocket expenses have been taken into account.

35. In line with this, in other datasets, median and mean time for developing STDR was 3 and 4 years. Hence, one possible reason why this study shows a low rate of detection of STDR was the short follow-up time. Even while long for a trial, it may be too short for the outcomes of interest.

STDR detection was a secondary outcome - the trial was designed for attendance. The low rates of STDR detection are in line with our current experience - 1.7% in the control arm is the same as in 2016/17. The primary outcome was attendance. As nobody within this study design would have been allocated to a screening interval longer than two years outcome at 3 or 4 years was not required for this RCT. A longer follow-up would have been ideal because of the low STDR rate but a 2 year follow up within the trial was the only feasible solution.

36. It is described that the attendance of participants during 24 months is used as a measure for safety. However, in the outcome section (Page 3), the authors state just considering the attendance at the first follow-up. These reviewers would suggest considering the attendance for the whole trial is more accurate.

We welcome this comment. An investigation on the equivalence in attendance rates between the two arms over a period of 24 months has now been added as a post hoc analysis. The interpretation does not change and equivalence is confirmed at the 5% equivalence margin. We now refer to this analysis in the manuscript [lines 271-276] and Supplementary Table 5.

37. A final major problem with the current manuscript is the use of a single risk prediction model, which was only internally validated, together with a single cut-off value, based on expert/stakeholder opinion**.**

The next stage of development of screening is dependent principally on the uncertainty about safety of varying intervals. This was the position prior to the findings in our RCT and our motivation in the design of our approach, introducing risk estimation and variable intervals. We recognise that this is only one model. The main aim was to test an effect on attendance of our new approach.

38. Lack of external validation limits the generalizability.

We agree with the reviewer and plan external validation but this was not within the scope of the award that funded this RCT. We recognise this as a limitation to the implementation of the findings of our study and have commented on this in the original submission. However we have shown that extended intervals In Liverpool are feasible. We appreciate that the safety and cost effectiveness depend not only on the population, the ability to obtain current risk factors and the risk model but also on the robustness of the retinopathy grading, which varies widely from programme to programme.

Also it is not clear why a new prediction model had to be developed, if quite some prediction models already exist, some of which have been externally validated.

The RCT is part of a 6 year NIHR Programme Grant for Applied Research. The programme grant started in February 2013. There are 7 workstreams which include a literature review; a prospective observational study of risk factors for progression to sight threatening diabetic retinopathy, treatment and visual impairment; development and internal validation of a risk calculation engine; health economics; RCT; exploring perceptions of screening and variable screening intervals amongst people with diabetes and professionals; knowledge transfer and preparation for implementation. The RCE needed to be developed and tested before the RCT could start and at that point although some risk models did exist at the start of the study no other such engine existed. The plan was always to develop a risk engine based on the local population.

39. The cut-off value could a priori have been optimized; balancing the trade-off between cases missed and cost savings. Now it is unclear whether the best possible personalized strategy has been followed.

Thank you for your comment. Our analysis examined the efficacy and effectiveness of the risk calculation engine's assignment to 6/12/24 month screening intervals applied to the total population at a two year time horizon (within trial). We agree that balancing the trade-off between cases missed and their cost savings would be an informative analysis, rather than simply observing a trade-off. We plan to undertake this as part of our future cost effectiveness work using a risk-based model.

Minor comments

40. Abstract: The background section mentions empowerment of persons with diabetes. However this statement is not addressed in the remainder of the paper. Suggestion is to explain in the discussion how personalized/risk based retinopathy screening contributes to such empowerment.

We have reflected on the term empowerment and changed in the summary to engagement. In response to Reviewer 1 comment no 1 we have added to the discussion a statement on the potential benefits to individuals from individualised screening and added a discussion point on better engagement with screening and overall diabetes care, "Further, in the era of personalised care a shortened interval in high risk people may increase focus on risk factor control and improve engagement with screening." [Lines 412-413].

41. Evidence and Introduction: The only existing and externally validated risk calculation engine mentioned is that of Aspelund et al. However, several more published prediction models exist and a couple of them have also been externally validated:

> Scanlon PH, Aldington SJ, Leal J, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess 2015;19(74):1-116. doi: 10.3310/hta19740
> Basu S, Sussman JB, Berkowitz SA, et al. Development and validation of Risk Equations for Complications Of type 2 Diabetes (RECODe) using individual participant data from randomised trials. Lancet Diabetes Endocrinol 2017;5(10):788-98. doi: 10.1016/S2213-8587(17)30221-8
> Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: cohort study. BMJ 2015;351:h5441. doi: 10.1136/bmj.h5441

We are aware of a number of risk prediction models in diabetes some of which include retinopathy. We reviewed them, including those by Hippesley-Cox, and Stratton and Scanlon, in our manuscript Eleuteri et al Diabetologia 2017 (ref 18) in which we described the risk calculation engine developed for ISDR. We have been clear in that paper and also in this current manuscript that further work on upscaling our risk based variable interval approach will be needed.

We are grateful to the reviewer for directing us to the RECODe publications, which appeared recently, and have referenced their most recent paper published in Diabetes Care in 2018. The RECODe equations are derived from a phase 3 trial cohort rather than a population cohort and have been externally validated. The aim appears to be to produce a clinical management tool rather than to be applicable to screening or extended intervals. The authors acknowledge that the retinopathy outcome is imprecise and that further validation in a longitudinal cohort is needed.

We realised that the evidence for extending intervals for low risk people from the Scanlon group comes from the Stratton paper which we have already referenced at ref 11 and so think that citing the Scanlon HTA report can reasonably be dropped to keep within the limit of 30 references.

We agree that the paper by Hippsley-Cox and Coupland is an important paper in which they described a similar method to predict 10 year rates of amputation and blindness. The relevance to our paper is not as strong as Aspelund et al or the work of the Scanlon/Stratton group as the only retinopathy end point they used was blindness and this was not measured in a validated method. We would be happy to include if the editorial team allow the inclusion of another reference.

42. The introduction seems to suggest truly personalized screening and use of the risk prediction model at each next visit to assess the best screening interval. This would e.g. imply that for some individuals the policy could be a screening frequency of 6-12-6 months, when their risk factors would vary over time. This is also suggested by the Procedures section.

However, the current manuscript reports on risk groups characterized by their screening frequency (at 6 months, 12 months and 24 months group, being determined from baseline risk assessment). This is inconsistent and confusing.

We have revised the methods section to clarify that the risk calculation is applied at each visit [lines 159-160] and added to discussion [lines 362-364].

43. Could the authors inform about the mean screening intervals in each of the risk groups? How often would individuals switch screening interval during study follow-up? This is informative, and informs about the need for truly personalized and adaptable screening, at least over a two year time horizon.

We thank this reviewer for suggesting this additional analysis. We have extracted the data on switching between screen intervals in the individualised arm over the 24 months of the trial. We have added in the manuscript the mean number of attended appointments per person over 24 months stratified by baseline risk [lines 270-274] shown in the table below. We have added a new Supplementary Table 6 to show the data on switching between intervals during the 24 months of the trial. 59% of people allocated to the medium risk (annual) group at baseline changed interval at least once in the 24 months, 75% of them to 24 months, adding further evidence of the potential reduction of burden of appointments for people in fixed interval screening. [Lines 274-276 in the results and 362-364 in the discussion].

Table: Breakdown of average number of attended appointments over 24m and at least one interval change over 24m per individual on the individualised arm by risk group.

Risk Group	n^	Expected follow-ups (n)	Attended follow-up appointments (n)	Mean number of attended appointments per person	Appointments at which the allocation changed (n)	Mean number of allocation changes per person	At least one change in allocation (n)	Percentage of individuals with at least one change in allocation
High	160	373	292	1.83	54	0.34	48	30.0
Medium	200	248	211	1.06	129	0.65	118	59.0
Low	1694	1694	1441	0.85	142	0.08	142	8.4

^Number of individuals with a recorded attendance over 24m outcome.

To fully interpret the implications of these data will require further modelling to correct for the censoring effect at the end of the 24 month period and the differential effect of changing risk factors and grading. We will add this to our post implementation monitoring programme and will report in the future.

44. Please refer to the risk groups consistently as high, medium and low risk and not as 6, 12 and 24 months, since that is too suggestive of constant intervals.

Thank you for this helpful suggestion. We have now made this clear throughout. Thank you.

45. Methods: The chosen primary outcome measure (equality of attendance rates) is mostly a process indicator of safety (persons not missing visits is a first requirement). The next outcome, cases of STDR missed seems much more relevant. But the study was powered on the primary outcome of course.

Attendance is a key outcome in diabetic retinopathy screening programmes since evidence shows that people who do not attend have risk factors that make them more vulnerable to diabetic retinopathy. For this reason we defined attendance as the primary outcome. Nevertheless, we considered STDR detection as the first secondary outcome. In addition, the low expected rates of STDR meant that for a study to be properly powered to assess non-inferiority in STDR detection, a sample size of about 20,000 would have been required, which was not feasible.

46. It is not clear why DR scores are not reported at all, and why baseline DR scores do not enter into the prediction model.

Baseline retinopathy scores are given in Table 1. In the Procedures part of the Methods section we state " Individualised arm participants were allocated by the RCE to a 6, 12 or 24 month interval at baseline and at each follow up visit using the most recent and clinical risk factor data". We have changed this to "At each screening visit individualised arm participants were allocated by the RCE to high, medium or low risk using the most recent screening and clinical risk factor data. The acceptable risk threshold was set at 2·5%, as agreed with the PPI group. The interval could vary at each follow-up visit depending on retinopathy levels and risk factors." [Lines 159-161].

Please note that baseline DR scores are used in the model. The risk model is based on a continuous-time Markov process. The patient state at each time point was defined by level of retinopathy and the risks for each transition were entered into the model (Eleuteri et al Diabetologia 2017 Ref 18).

47. The chosen variable selection methods (AIC) is not fully clear. What was the total set of variables available? How and in what order were different models compared?

We included this information in our publication Eleuteri et al (ref 18). In addition, the electronic supplementary material developed for that publication provides a detailed description of the candidate covariates, covariate selection and model fitting.

48. The Procedures state that the most recent visit of a participant was used to allocate individuals to 6, 12 or 24 months intervals. Not only at baseline, but also at each follow-up visit. However the results are described by risk group, classified as 6 m, 12 m or 24 m intervals screening. This is not clear. What risk group do individuals belong who first classify as 6 and then as 12 or any other combination? This is more clearly described in the statistical analysis part. The risk groups are determined by RCE score at baseline. It may be advisable to alter the names of the risk groups, since it is confusing.

We have changed to high, medium and low risk.

49. In the ITT analyses, what was the assumption concerning persons with unknown STDR status? Similarly for other endpoints? What was assumed in the ITT?

People with unknown STDR status did not contribute to the main ITT analysis to avoid the adoption of any assumptions about their status. Nonetheless, and as indicated in the manuscript, ITT analyses that included every participant according to assigned randomisation were also conducted for sensitivity

analyses. Multiple imputations generated using generalised linear models (GLM) dependent on the baseline characteristics (PROC MI, SAS) assessed the effect of missing values on both PP and ITT datasets. We also conducted multiple imputations under the assumption of the most conservative and less conservative scenarios. These extra analyses, which were described in the Statistical Analysis Plan, are now included in the manuscript (as a supplementary table).

50. It is a bit strange that only a limited sample completed the QOL questionnaires. Their burden is not large (EQ5D 5L only contains 6 questions) and since the study was not powered to differences in QOL, using an even smaller sample for this seems not justified.

We set our study originally to recruit around 700 cases although we over-recruited. We agree that the EQ5D is not burdensome in itself but we were concerned that the accompanying HUI3 could have an impact on the smooth running of the clinics. We have added the statement in the discussion [lines 407-408], "… our limited sample size was chosen to minimise disruption in the screening clinic." We did test the data using multiple imputation methods.

51. Were levels of risk factors like blood pressure included in the RCE as such, or was the level of medication also included? **Systolic blood pressure was used but not medication. The detailed development of the RCE has already been published. Eleuteri A, et al. Diabetologia 2017 ref 18**

52. How was STDR defined? This is in the protocol, but for an important outcome like this it is more clear to shortly describe the definition in the methods. **This has now been defined at lines 164-165.**

53. Results: The reviewers suggest to keep to the names of low, middle and high risk groups, since 6, 12 and 24 months is confusing. As these were only the baseline frequencies of these groups. **We have changed this for clarification. Thank you.**

54. The event rates for STDR were quite low. It is not clear how CI for differences in such low percentages were estimated. This should take into account the small proportion.

We are grateful for this comment. As explained in our response above (editorial comment 5-p) we now report CIs based on the Wilson score approach for the non-inferiority study. As expected the CIs remain very similar.

55. DR is here also reported as an outcome but was not described in the methods very clearly.

We are now reporting DR, together with other secondary outcomes, in supplementary tables.

56. Discussion: It is surprising that the RCE did not include pre-existing DR as a predictor. Was this tested? Why not?

Please see answer to comment 46. Pre-existing DR was taken into account in the RCE. In lines 152-153 we state that the RCE is a Markov multi-state model, where the states are defined according to the retinopathy levels in both eyes, and state transitions depend on risk factors. DR at screening visit is the base risk. Pre-existing DR is included through the population transitions introduced into the multi-state model.

57. Why were intervals longer than 24 months not included?

We would have liked to have included a longer interval. In the Hoorn Study some people have 4 or 5 year intervals but are well failsafed as they are seen in the diabetes centre. In the previous paper from members of our group (Younis et al, Lancet 2003) we estimated an interval for people with no DR to have a 95% chance of not developing STDR to be 5.4 years. We recommended that a 3 year interval would be reasonable. However the constraints of the funding, stakeholder resistance, sample size and the trial design prevented us from extending beyond 24 months. This constraint does beg the question whether a 3 year interval should be considered. There is a sentence in the discussion on this. 2 yearly screening for low risk groups has already been proposed in the UK. [Line 390].

58. Ref 26 concerns a Dutch not a Danish population. **Changed. Line 367. Thank you!**

59. Refs 11 and 28 are only a couple of existing publications on risk models in retinopathy. See the references above among others.

We agree and would have liked to have covered more publications, but we are limited by the number of references allowed.

60. It might be relevant to discuss how many months a screening might be performed to still be "in time" to detect STDR. That is, what is the damage when STDR is discovered at 12 months rather than at 6 months?

Also what happens to patients when STDR is detected? What are the costs and quality of life consequences of being referred for STDR?

Unfortunately we feel these interesting points our outside the scope of the trial. Our time horizon for reporting outcomes and the CEA was restricted to the 2 years of the RCT. However we intend to undertake some additional modelling and some limited data collection including over a longer time horizon which will allow us to investigate the outcomes post STDR detection.

61. Tables are inconsistent in terms of precision of the numbers reported, as reflected by the number of decimals/figures. Especially table 5. **Corrected**

62. For table 1 it would be informative to provide information about what variables were significantly different (e.g. at 5% level). E.g. for retinopathy level this might be the case?

Thanks for this suggestion, Table 1 has been revisited and it now shows which variables were significantly different across the risk groups. As per current guidance we have not compared control and individualised arms directly.

63. A table with DR scores would be informative.

This data is in Table 1

64. It is not informative to report negative ICERs, since these could reflect cost savings at positive health benefits, as well as additional costs at health losses. Rather it is preferable to state "dominant" or "dominated" in this case**.**

We appreciate the problem of interpreting the negative ICER. Our new summary Table 3 now incudes them, and we better address this problem by reporting incremental net monetary benefit as the statistic of choice for the health economics.

65. Figures and tables with incremental costs in relation to cases of STDR missed would be informative.

THELANCET-D-19-02765_response Broadbent et al.

We appreciate the problem of interpreting the negative ICER as specified in our new table 3. We address this in a number of ways: in the footnote of table 3 we discuss where in the cost effectiveness plane these values lie dependent on a positive or negative value, while keeping the ICERS inset in the table due to their informative nature for those familiar with cost effectiveness analysis. We also discuss the interpretation of these ICERs within the results and discussion section of the health economic supplement. In practical terms we have also reported the incremental net monetary benefit alongside ICERs which are not confounded in their interpretation. While we agree in such a case it should be reported dominant or dominated; since dominance is only found in costing differentials (and stated as such in the text) and our differences in QALYs cross the zero threshold we cannot claim strict dominance of individualised versus annual screening.

Reviewer #6

66. The primary outcome showed equivalence in attendance rates between the two arms the first visit. The higher loss to follow up in the individualized arm is of concern particularly since the Liverpool region has a well- developed screening program with high participation rates.

We would like to point out that a higher loss to follow up for the primary outcome in the individualised arm was expected, given that the majority of participants in the that arm were allocated to the 24 month screening interval, allowing therefore for a higher chance of being loss to follow up. Contrary, loss to follow up over a constant period of 24 months in both arms show similar rates. The manuscript includes a sentence referring to this point. "Loss to follow up was higher in the individualised arm, 101 (4·5%) compared to 41 (1·8%) in the control arm, largely due to the longer follow-up period of 24 months experienced by low risk participants (83·4% of the individualised arm) compared to the 12 months follow-up in the annual arm".

The chances of even greater losses to follow up in low risk participants may be even higher in other regions with less organized facilities. **This point has also been picked up by reviewer 4. The Liverpool programme is well established with relatively low rates of disease. We have been careful to emphasise this in our interpretation of our findings - that introduction into other programmes should be conducted with careful monitoring.**

It would be helpful to know what efforts were made by the research team to establish the reasons for loss to follow up? **This is covered in the additional supplementary tables.**

In this context the proportions within the two groups experiencing serious adverse events during their follow up may provide some answers. Persons with DM are more likely to have conditions such as MI, stroke and the risk of dying is also higher. So it would be very helpful to know how these events were distributed between the randomized arms and if these rates were different (e.g. Lower detection rate in the individualized group may suggest that co morbidities may have prevented attendance).

We discussed SAEs at the design phase and determined that as the RCT was not a CTIMP we would not collect SAEs or for development of ATEs.

67. It is important that a clear definition of screen positive is provided.

Thank you. Added at lines 146-147.

68. I did not find the following data which is also important. How did the clinical and ocular characteristics of those persons who declined to take part in the study differ from those who agreed?

We looked at the summary measures of baseline characteristics (demographic, clinical and ocular) in the two groups: consented/randomised patients and patients who did not consent. While the data collected in the consented/randomised group is of good quality (as expected since this is trial data), for the non-consent group we had to rely on the data available in EMIS and OptoMize, and for some variables, such as ethnicity and smoking status, the missing rate was high. Nevertheless, we could not identify clinically significant differences in level of retinopathy, total cholesterol, diastolic or systolic blood pressure and HbA1c. We however observed than people who consented tended to be younger (63 vs 67 in the non-consent group) and male (60% vs 54% in the non-consent group).

Reviewer #7

69. It is understandable that attendance to screening should be an important outcome. However I wonder if the trial would be more valuable if it could confirm attendance at subsequent visits, perhaps to a second visit, or even more useful, to consider detection of STDR as the primary outcome. Why the authors did not choose detection of STDR as primary outcome?

These are interesting points which we did consider during the study development phase. The main issue preventing adoption of extended intervals in screening is the concern about non-attendance, a point that we have covered in answers to other reviewers (comments 35, 37, 45) and within the original submission.

70. There are insufficient details on the categorisation of risk to sight threatening diabetic retinopathy. Did the investigators develop a risk calculator (and was it validated) or have they used a known and validated risk calculator?

This is described in the procedures subsection of the methods. Further detail is available in Eleuteri et al Diabetologia ref 18.

71. It is unclear whether during the RCT the risk profile could change as diabetic-control relevant data was updated bi-monthly. Can the authors explain?

We have addressed this in response to other reviewers (comment 42, 43, 53). Our individualised arm comprised variable intervals based on risk. So at each visit the interval could change. We apologise for the ambiguous text in lines 159-161 and have revised for better clarity. We have added also added a new supplementary table 6.

72. A secondary outcome was detection of STDR, with a non-inferiority margin of 1.5%. It seems that the sample size gave the study 60-65% power to test non-inferiority in detecting STDR. This is a relatively weak power to be recommend a change of current policy in the English NHS.

We agree that our analysis on STDR detection has relatively weak power. This is mainly due to the low event rate for STDR development. We have referred to this analysis as secondary. On its own it is insufficient to recommend a change in current policy, but taken in conjunction with our primary analysis, the other secondary analyses and our CEA we believe that our recommendation on policy change are reasonable.

73. The authors stated that detection rates of STDR were non inferior. However, the authors also stated that "for the high and medium risk groups... non-inferiority was not confirmed for these two groups, which may be explained by the uncertainty of the estimates due to small numbers." The high and medium risk groups are potentially the most concerning (to clinicians and patients).

We agree with this remark. We would like to highlight that although non-inferiority was not confirmed for the high risk group, the detection rate was higher in the individualised arm, (13.4% vs 11.4% in the annual arm) which is desirable and relates to the fact that this group is seen more often in the individualised arm. For the medium risk group, the two arms used the same screen interval of 12 months, and consequently the difference in STDR detection rates was relatively small. One of the key features of individualised screening is that the majority of people are low risk (which is important for its cost-effectiveness), but this inevitably means low numbers for the medium and high risk groups which, as we know, results in larger CIs.

74. The authors' conclusion is that individualised risk-based intervals can be safely introduced. However, I do not think the conclusion is justified by their data.

Following on from our response to comment 72 we acknowledge that there are limitations to our findings which we have tried to convey in the discussion, research in context, and concluding remarks. Other reviewers have been supportive. We do believe that our study provides convincing evidence on the safety of extending intervals with the important caveats that further scale-up is required with appropriate monitoring and the findings should only be applied to settings such as ours. We believe that established programmes with access to clinical and screening information could introduce individualised risk-based intervals as long as there was good failsafe operating and if the attendance rate was at least as good at that in Liverpool.