**Personalized risk-based screening for diabetic retinopathy: a multivariate approach vs. the use of stratification rules**

*Running title:* Screening for DR using longitudinal data

Marta García-Fiñana,1 David M Hughes,1 Christopher P Cheyne,1 Deborah M Broadbent,2,3 Amu Wang,2 Arnošt Komárek,4 Irene M Stratton,5 Mehrdad Mobayen-Rahni,2,6 Ayesh Alshukri,2 Jiten P Vora,7 Simon P Harding.2,3

1. Department of Biostatistics, Institute of Translational Medicine, University of Liverpool, Liverpool, UK

2. Department of Eye and Vision Science, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK

3. St. Paul’s Eye Unit, Royal Liverpool University Hospital, Liverpool, UK

4. Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

5. Gloucestershire Retinal Research Group, Cheltenham General Hospital, UK

6. Department of Medical Physics and Clinical Engineering, Royal Liverpool University Hospital, Liverpool, UK

7. Diabetes and Endocrinology, Royal Liverpool University Hospital, Liverpool, UK

Corresponding author:

Dr. Marta García-Fiñana

Reader in Biostatistics

Department of Biostatistics, Institute of Translational Medicine

Block F, Waterhouse Bld

1-5 Brownlow Street

Liverpool, L69 3GL

Tel 0151 794 9755

martaf@liverpool.ac.uk

**Word count:**3,950

1 Table and 3 figures

2 Supplementary tables and 1 supplementary figure

**Personalized risk-based screening for diabetic retinopathy: a multivariate approach vs. the use of stratification rules**

**ABSTRACT**

AIMS - Timely detection and treatment of sight threatening diabetic retinopathy (STDR) is key for the avoidance of visual impairment. This study proposes a multivariate approach to identify patients who will develop STDR within a one-year screen interval, and explores the impact of simple stratification rules on prediction.

MATERIAL AND METHODS – A 7-year dataset (2009-2016) from people with diabetes (PWD) was analyzed using a novel multivariate longitudinal discriminant approach. Level of diabetic retinopathy assessed from routine digital screening photographs of both eyes was jointly modeled using clinical data collected over time. Simple stratification rules based on retinopathy level were also applied and compared with the multivariate discriminant approach.

RESULTS – Data from 13,103 PWD (49,520 screening episodes) were analyzed. The multivariate approach accurately predicted whether patients developed STDR or not within a year from the time of prediction in 84.0% of patients (95%CI: 80.4%-89.7%), compared to 56.7% (55.5%-58.0%) and 79.7% (78.8%-80.6%) achieved by the two stratification rules. While the stratification rules detected up to 95.2% (92.2-97.6) of the STDR cases (sensitivity) only 55.6% (54.5-56.7) of patients who did not develop STDR were correctly identified (specificity), compared to 85.4% (80.4%-89.7%) and 84.0% (80.7%-87.6%), respectively, achieved by the multivariate risk model.

CONCLUSIONS – Accurate prediction of progression to STDR in PWD can be achieved using a multivariate risk model whilst also maintaining desirable specificity. While simple stratification rules can achieve good levels of sensitivity, our study indicates that their lower specificity (high false positive rate) would therefore necessitate a greater frequency of eye examinations.

**INTRODUCTION**

Early detection and treatment of sight threatening diabetic retinopathy (STDR), a stage of diabetic retinopathy (DR) requiring referral to an ophthalmologist, is important to avoid visual impairment in people with diabetes (PWD) (1-7). Risk factors for the development and progression of DR have been identified in epidemiological and observational studies (8-21). Evidence of the influence of clinical variables in DR progression, such as duration of diabetes, glycosolated haemoglobin (HbA1c), type of diabetes, etc., has been widely reported (3, 8, 9, 19, 21-23).

Given the low annual incidence rate of STDR in the population with diabetes in developed countries (<3%), personalized risk-based screening should offer a cost-effective approach to reduce the economic burden on health systems without compromising efficacy (8, 9, 24). Identifying STDR early is important, not just to enable prompt and effective treatment and thus maintenance of vision, but also to allow cost-effective screening intervals tailored to patients’ needs. The potential benefit of a personalized screening approach (where patients at higher risk are screened more often than those at low risk) depends on the accuracy and validity of the predictive model utilized and on the associated costs. Data collected from a patient over time (longitudinal data) that captures changes of clinical markers could be used to improve the accuracy of a predictive model. In this study, the individual trajectories of the clinical profiles in people with diabetes are used to develop and to validate a predictive model for STDR.

We jointly modeled demographic and clinical data to characterize the baseline level of retinopathy and its changes over time. We applied a risk-based longitudinal multivariate approach that enables the identification of patients who will develop STDR within a year from the time of prediction. In the UK and elsewhere, there is a current debate as to whether a simple rule based on the patient’s level of retinopathy alone or alternative multivariate clinical models should be used to determine risk-based screening intervals for STDR (21-27). Risk stratification for development of STDR based on just the results of two screening episodes has been proposed by Stratton *et al*. (25). They observed that the annual rate of progression to STDR was 0.7% for patients with no DR at two consecutive annual digital photographic screenings, 1.9% for patients with no DR in either eye at first screening but mild non-proliferative DR (mild NPDR) / background DR (BDR) in just one eye at second screening, and 11% for patients with mild NPDR / BDR in both eyes at both screenings. Here we compare the overall accuracy of our multivariate model with two *simple* risk stratification rules, including the rule by Stratton *et al*., 2013, which has been agreed by the UK National Screening Committee to be introduced in England within the next years (26).

**MATERIAL AND METHODS**

**Study participants and design**

Data from 13,103 people with diabetes mellitus registered with a general (family) practice (GP) in Liverpool were included in our model. Demographic and systemic risk factors data from primary care systems (EMISweb, EMIS Health Ltd., a company based in Leeds which supplies electronic patient record systems and software used in general practice in England) and level of retinopathy obtained from the Liverpool Diabetic Eye Screening Programme (49,520 screening episodes, from 2009 to 2016 with a median follow-up of 6 years; OptoMize, EMIS Health Ltd.) were linked in a purpose built data warehouse. Patients were offered annual screening for DR following national recommendations (28). When patients did not attend their first appointment for screening they were offered a second appointment (usually within 6 weeks of the first appointment). At screening appointments patients had at least two 45° digital retinal photographs taken per eye, according to national guidelines, which were graded by accredited technicians to assess the level of DR.

A data sharing agreement allowed access to GP data via the Liverpool Clinical Commissioning Group (CCG). Practices were approached between 2013 and 2016 and all 92 within the Liverpool area agreed to participate. Patient consent was sought via an opt-out approach approved by the local research ethics committee (13/NW/0196). Data from PWD who had opted out of the study (7.3%) were not considered for the analyses. Patients with only one clinic visit (8.5%) or with no recorded clinical visit within a time window of 18 months before the final visit (8.7%) were also excluded from the analysis.

**Study variables**

Available demographic and systemic risk factors data included age, sex, ethnicity, recorded diagnosis of diabetes (time and type), attendance for screening, HbA1c, diastolic and systolic blood pressure (DBP, SBP), total cholesterol and eGFR recorded over time during follow-up. The following disease states were considered for each eye: 1. No DR detected; 2. Non-referable DR (mild NPDR / BDR) and 3. STDR defined as moderate/severe pre-proliferative DR or proliferative DR and/or maculopathy, i.e., any of the following features: multiple blot haemorrhages, venous beading, intraretinal microvascular abnormalities, new vessels, preretinal/vitreous haemorrhage, fibrovascular proliferation, exudates within 1 disc diameter (1500 μm) of the foveal centre, group of exudates within the macula more than ½ disc area in size, or retinal thickening within 1 disc diameter of the foveal centre.

Patients with STDR at the start of the prediction period were excluded from the analysis. For the purpose of this analysis, the values of the time-dependent clinical variables closest to the time of the screen episodes (i.e., annual screening episodes) were used. A complete case analysis, similar to Scanlon *et al*., 2015 (9), was followed and screening visits for which model covariates were not available were excluded from the final model.

**Model development and statistical analysis**

We have recently developed a multivariate discriminant approach, which can be used to predict the future status of a patient using their clinical history (29,30). Here we applied this statistical approach to estimate the risk that a patient will develop STDR in either/both eyes within a one year period, and this is achieved by using the demographic and longitudinal primary care data to jointly model the changes in level of retinopathy over time for both eyes (see Figure 1, supplementary material). Our approach is based on the following rationale: two longitudinal models are generated using part of the dataset (training dataset), one for each of two possible prognostic groups (patients who develop STDR and patients who do not develop STDR within a year). These two models focus on modelling the progression from no DR to mild NPDR / BDR, making use of the biochemical and demographic records of the patients. The status of a new patient is then predicted depending on which of the two models the new patient’s clinical profile is statistically *closest* to (29). This statistical approach calculates the risk of a new patient developing STDR within a year from the time of prediction, and this risk can be updated each time new data become available for the patient.

The transition from no DR to mild NPDR / BDR was modeled using a bivariate generalized linear mixed-effects model that takes into account the correlation between measurements at different time points for the same patient. It is bivariate since it captures the measurements from both the right and left eye in a single model. The grading in each eye was considered as a binary longitudinal variable with 0 representing no DR and 1 denoting mild NPDR / BDR. Correlation between repeated measurements for a patient and between retinopathy grading in each eye was modeled using random (patient specific) intercept terms in the mixed model. A 2-component mixture of Gaussian distributions was specified to allow flexible modelling of the joint distribution of random effects. Longitudinal models included clinical risk factors that influence the changes in retinopathy level over time. Due to the complexity of the statistical model, the Markov Chain Monte Carlo (MCMC) method was used to estimate model parameters (29, 30).

Penalised expected deviance (PED) alongside a forward selection approach, in combination with clinical judgment, was applied to identify the relevant demographic and clinical risk factors that influence changes in retinopathy level over time. Models were compared using PED, which penalizes for model complexity and is suitable for complex hierarchical models (31). Due to the stochastic nature of MCMC, it is possible that different random starting values (seeds) generate slightly different models. Hence, in order to check the stability of the model, the process was generated for multiple seeds. Two training datasets, involving data from 70% of patients in each of the two prognostic groups, were used to build the model and data from the remaining 30% were used to test the predictive accuracy of the model. Training and test sets were randomly generated 100 times and the results were averaged. The statistical analyses were performed in R version 3.0.2 using the package mixAK (32).

To allow for the fact that patients have been observed for different lengths of follow up period, time since first screening was included as a covariate in the longitudinal models. All patients had been followed-up for a minimum of two years. To develop the models we considered the period of time from the start of their observations up until the point 1 year before their final visit (in order to be able to predict the clinical status 1 year after). For patients who developed STDR, the final visit was defined as the time at which STDR was detected (i.e., data beyond STDR detection is ignored). For patients who did not develop STDR the final visit was the last recorded visit. Therefore, group memberships (i.e., whether or not the patient developed STDR within one year of the prediction visit) were known for all patients.

The fitted mixed models, one for each prognostic group, were used in a longitudinal discriminant analysis to predict the likelihood that a new patient will/will not develop STDR within a year. In particular, the likelihood of the new patient’s data coming from each of the two mixed models was assessed and then weighted by the prevalence of each group to give a probability of developing STDR within a year. If this probability was greater than a threshold (chosen through analysis of a Receiver Operator Characteristic (ROC) curve), then the patient was classified as developing STDR within a year and otherwise they were classified as non-STDR.

The two prognostic models were subsequently used to predict for a new patient (test set) the likelihood of developing/not-developing STDR within a year. Intuitively, the patient is linked to the group with the model the new patient’s profile is *closer* to. The level of accuracy of the multivariate approach was assessed using the area under the ROC curve and its 95% confidence interval (CI). We also assessed the values of sensitivity (percentage of patients, out of patients who truly developed STDR, who were correctly identified by the model), specificity (percentage of patients, out of patients who did not develop STDR, who were correctly identified by the model) and the probability of correct classification (percentage of patients correctly classified), with their corresponding 95% CIs. The threshold chosen was associated with the point on the ROC curve nearest to the top left corner (i.e. it provides the best balance in terms of number of patients correctly identified as not developing STDR and those correctly identified as developing STDR).

*Simple stratification rules*

Annual screening for DR has been adopted by several national screening programs (e.g., 27, 28, 33). Risk-based stratified screening intervals are likely to be introduced in a number of countries within the next years in order to cope with the imminent significant rise in the number of PWD. We explored with our data the overall accuracy of prediction using simple stratification rules based on retinopathy level alone to identify low and high-risk patients (the latter group consisting of patients who develop STDR within a year). Ideally, low-risk patients could be offered 2-year screening intervals or longer intervals. We define *sensitivity* as the percentage of patients, out of the patients who develop STDR within a year, who are correctly predicted by the rule (and therefore allocated to annual screening intervals). Specificity is defined as the percentage of patients, out of the patients who do not develop STDR within a year, who are correctly predicted by the rule as not developing STDR within a year (and therefore are allocated to biennial screening intervals). We also calculated the reduction in the number of screening episodes achieved by simple stratification rules when compared to the currently recommended annual screening.

Subgroup analyses were conducted to further explore the effect of diabetes type on the classification performance.

**RESULTS**

7-year data from 13,103 people with diabetes were included in our model. The median follow-up was 6 years between March 2009 and January 2016. The demographic and clinical characteristics by prognostic group are provided in Table 1. Screening visits for which model covariates were not available were excluded from the final model (12%). Compared to the non-STDR group, we found that patients who developed STDR during the follow-up period were more likely to be men, younger, have type I diabetes, longer disease duration, higher HbA1c and were more likely to have missed screening appointments.

As expected, the majority of patients who developed STDR (74%) within a year screen interval exhibited mild NPDR / BDR in both eyes at their previous screen visit (prediction visit); only a small percentage of patients (12%) who developed STDR showed no DR in either eye at their previous screen visit. This trend is reversed for patients who did not develop STDR (78% showed no DR in both eyes and only 9% showed mild NPDR / BDR in both eyes at the time of prediction).

**Table 1**

As expected, the risk factors for the progression from no DR to mild NPDR / BDR show similar ORs for the right and left eyes with similar interpretation (data from both eyes were jointly modeled and the full model specification is given in Table 1 of supplementary material). For simplicity, we report one OR for each risk factor. Ethnicity and eGFR were not included in the model due to the high rates of missing values observed (18% and 36%, respectively) and the lack of ethnic representation (predominantly white).

*Multivariate model for patients who developed STDR*

For patients who developed STDR within a year, the multivariate model showed that progression from no DR to mild NPDR / BDR during the follow-up period prior to STDR was associated with diabetes duration with an odds ratio (OR) (per 5 years of disease duration) of 1.78 (95%CI 1.21-2.62) and with their previous screening appointment missing with an OR of 2.12 (1.05-4.42). Other factors such as sex, age, HbA1c, systolic and diastolic blood pressure, diabetes type and cholesterol level were not found to be statistically significant. The variable HbA1c was included in the final model for completion (despite its lack of significance in this group) due to its relevance in the literature. The inclusion of HbA1c nonetheless did not affect the clinical interpretation of the other coefficients in the model.

*Multivariate model for patients who did not develop STDR*

For patients who did not develop STDR, progression from no DR to mild NPDR / BDR was associated with diabetes duration with an OR (per 5 years disease duration) of 2.25 (95%CI: 2.10-2.40) and type I diabetes with an OR of 2.44 (1.88-3.21). Also, in this group the transition from no DR to mild NPDR / BDR was less likely to occur as time progressed (OR=0.97 per year, 0.95-0.99) as opposed to the STDR group, for whom the transition from no DR to mild NPDR / BDR was more likely to occur as time increased (OR=1.60 per year, 1.33-1.93). SBP and HbA1c showed statistically significant associations although the ORs showed a lower effect compared to the risk factors listed above (for SBP every 10 mm/Hg, OR: 1.07 (1.04-1.10); and for HbA1c every 10mmol/mol, OR: 1.04 (1.01-1.07)).

The random intercepts of the model account for the within-subject variation and encapsulate the underlying “state of health” of the patient not explained by the observable covariates. Within the group who developed STDR, there were two subgroups of patients; a group of just over a third of patients who had a high initial risk of progressing from no DR to mild NPDR / BDR (weight=29.2% in Table 1 of the supplementary material) and a remaining group with a lower initial risk of progression (although in both cases patients initial risk of progression was much higher than in the No STDR group). In the group of patients who did not develop STDR, just over a third of the patients belonged to a group with a very low initial risk of progression from no DR to mild NPDR / BDR (weight=36.2%) with the remaining patients having a higher initial risk. The model takes into account the correlation between the right and left eye outcomes through the covariance matrices.

**Figure 1**

*Accuracy of the multivariate discriminant tool*

The level of accuracy of the multivariate approach shown by Figure 1 indicates that 85.4% of patients who developed STDR within a year were correctly identified by the model (sensitivity, 95%CI: 80.4-89.7) and 84.0% of patients who did not were correctly identified by the model (specificity, 80.7-87.6). The area under the ROC curve (ROC AUC) was 0.90 (0.86-0.92). The probability of correct classification was 84.0% (80.4%-89.7%). Figure 2 shows the percentage allocated to each prognostic group based on the level of retinopathy at the time of prediction. Only 1.8% of patients who showed no DR at the time of prediction were predicted to develop STDR within a year (i.e., the predicted risk was greater than the selected cut-off for 2% of patients in this group). This percentage increased to 57.3% for patients who had mild NPDR / BDR in one eye only, and to 97.3% for patients who had mild NPDR / BDR in both eyes at the time of prediction.

**Figure 2**

The effect of diabetes type on the classification performance was further explored by conducting subgroup analyses. For patients with type 1 diabetes, our model achieved a ROC AUC of 0.81 with optimal sensitivity of 84.3% and 78.0% specificity. For patients with Type 2 diabetes, the ROC AUC was 0.90, with optimal sensitivity of 85.6% and specificity of 84.3%.

*Comparison to simple stratification rules*

When we assessed the overall predictive accuracy of the simple stratification rules to identify low and high-risk patients, we observed that if we apply the simple rule that patients with no retinopathy in either eye during a two-year period with two successive annual screening episodes are predicted as not developing STDR within a year, and otherwise as developing STDR (25, 26), the values of sensitivity and specificity were 95.2% (92.2, 97.6) and 55.7% (54.5, 57.0), respectively (Figure 3, left panel). Alternatively, if only the level of retinopathy of the screening episode at the time of prediction is considered, and patients are predicted as not developing STDR if neither of the eyes show DR and as developing STDR otherwise (27), sensitivity and specificity are 87.5% (82.4, 92.7) and 79.5% (78.6, 80.4), respectively (see Table 2, supplementary material). While the sensitivity drops by about 8% with the second rule compared to the first rule, the specificity increases by 24%, and consequently we observed a significant reduction in the number of screening episodes required (from a reduction of 27% with the first rule to a reduction of 39% with the second rule). The probabilities of correct classification were 56.7% (55.5%-58.0%) and 79.7% (78.8%-80.6%) for the first and second stratification rules, respectively.

Statistical comparisons in sensitivity and specificity between the multivariate approach and the two simple stratification rules demonstrated that while the level of sensitivity of the first rule is significantly higher when compared to the multivariate approach (9.8% difference, (4.9,14.7)), the specificity of the multivariate approach was significantly higher when compared to the first rule (28.3% difference, (24.9, 32.3)). The multivariate approach also showed a significantly higher level of specificity when compared to the second rule (4.5% difference, (0.8, 8.2)), while the increment in sensitivity by the second rule was not statistically significant at the 95% confidence level (2.1% difference, (-0.2, 7.4)).

**Figure 3**

**DISCUSSION**

We are heading towards personalized medicine where patient management can be tailored based on individual risk of disease or response to treatment. In particular, the predicted risk of developing STDR can be used to recommend personalized screening intervals. Annual screening is currently adopted in many national screening programmes at a considerable cost to health services (9). Longer screening intervals for those at low risk of developing STDR have been already introduced, for example in Iceland using a model that accounts for the level of retinopathy as well as other clinical information, including type and duration of diabetes, HbA1c and blood pressure (34-36). Additional models have been recently proposed in the literature to tailor screening intervals for DR (e.g., (8, 37). Longer screening intervals for patients with a low risk of developing STDR are expected to be implemented in England and Wales within the next years to cope with the increased economic burden triggered by the steady increase in disease prevalence and the lack of extra funding for screening. Risk-based approaches are timely in that they could be used to identify patients at high risk so that they can be closely monitored and treated earlier, while the majority of low risk patients can be screened less often allowing the optimization of limited health resources.

In this paper we report the results of a recently developed multivariate longitudinal approach (29, 30) to predict the risk that a given patient with diabetes will develop STDR within a year. The model shows high levels of classification accuracy (sensitivity and specificity were 85.4% and 84.0%, respectively). The AUC was 0.90, which is higher than the AUCs (0.79 and 0.76) previously reported in (9) and (35), respectively. A similar AUC (=0.90) was reported in (8), and although the specificity reported in (8) was 90%, the level of sensitivity was much lower than the sensitivity achieved with our multivariate model (67% versus 85.4%).

There are a number of advantages of our approach. From a methodological point of view, the approach is robust against misspecification of the distribution of the random effects term, which is a term that takes into account the correlation between measurements at different time points in the model (29). The approach has the potential to develop dynamic models where the risk can be recalculated every time new data from the patient become available (29, 30). All 92 practices approached within the Liverpool area agreed to participate which demonstrates the screening coverage data. Limitations of our study include the possible misclassification in level of retinopathy during grading, that the costs of misclassification were not considered (which differ between STDR and non-STDR misclassification) and that only internal validation was conducted. We acknowledge however, that a comprehensive assessment of the model’s predictive performance would require external validation using data from a different cohort. For example, the predictive accuracy of our model needs to be validated for different ethnic groups. Our dataset, which included predominantly white individuals, did not have representative samples across different ethnic groups. The differences in performance between our approach and the stratification rules must be explored using different cohorts to confirm the reproducibility of our findings.

Several multivariate regression models have been proposed over the years to predict DR risk (e.g. 9, 15, 16, 35, 36, 38, 39). In these models retinopathy grading and clinical variables are treated as predictor variables in the model. For example, Aspelund et al., 2011, proposed a risk algorithm for development of STDR based on the Weibull proportional hazard model (35). The predicted tool included type and duration of diabetes, HbA1c, blood pressure and presence of nonproliferative DR (defined as a binary variable: Yes/No). The recommended screening intervals generated by their risk model (which ranged from 6 months to 5 years) were estimated to achieve 59% fewer visits than with annual screening. This reduction is larger than the reduction achieved with our predictive model (41% reduction), which is due to the fact that screening intervals longer than 2 years were considered in their allocation. However, the accuracy achieved by their model was lower (e.g., 0.76 vs 0.90 for the AUC) and a significant limitation of their model is that it uses historical data on risk factors. Model coefficients in (35) were fitted separately for type 1 and type 2 diabetes, while type of diabetes was added here as a covariate in the longitudinal models. The subgroup analysis conducted by type of diabetes indicates that our classification tool is comparably sensitive irrespective of type of diabetes, with a small reduction in specificity in the less common Type 1 patients. We believe this provides confidence that the model can classify well regardless of the type of diabetes.

The conceptual idea behind our approach is different to previously proposed multivariate regression models. The profiles of level of retinopathy over time and prior to STDR development are first modeled using demographic and primary care data, and the risk that a new patient at a specific time point develops STDR is estimated based on these profiles (which are in turn driven by demographic and time-dependent primary care data). Markov chain models are based on a similar idea where the risk to move from one state to another state is modeled, and where state is defined by level of retinopathy (8, 37, 40). Rather than proposing an alternative predictive model to predict STDR in competition with those already developed, we devised an alternative model, with similar levels of accuracy that naturally describes the influence of clinical data (primary care/missing appointments data) on progression of diabetic retinopathy and how these relationships affect the risk of developing STDR.

The models discussed above focus on DR progression. The identification of DR in newly diagnosed individuals has been also subject of research. For example, in Cichosz et al. (41), a linear classification tool was proposed to predict mild or moderate non-proliferative diabetic retinopathy in newly diagnosed people with type 2 diabetes.

We acknowledge that alternative techniques such as neural networks could also be considered. Despite their advantages, such as an ability to detect nonlinear relationships between the outcome and dependent variables, their limitations include the inability for clinical interpretation (i.e., seen as a “black box”) and proneness to overfitting. The advantage of the method we have proposed is that it can deal with the complex structure of the data (i.e., longitudinal data collected over time, correlation between the measures from the right and left eye, outcome variables of different type) while allowing for clinical interpretation of the model.

In territories with established screening programmes, a clinical debate is taking place on whether known clinical factors for development and progression of DR (such as duration, HbA1c and type of diabetes, etc.) should be used in a decision model to tailor screening intervals for DR. Given that retinopathy level is the most informative factor on progression to STDR, a simple rule that makes use of this biomarker alone is being considered to determine risk-based DR screening intervals in the UK (26). The English Diabetic Eye Screening Programme has proposed an extension to 2-year intervals for patients with no retinopathy in either eye during a two-year period with two successive annual episodes. Our data suggest that if we apply this simple rule, for which patients with no retinopathy in either eye during a two-year period with two successive annual screening episodes are recommended biannual screening intervals, and annual screenings otherwise, a high level of sensitivity would be achieved (95%). This means that only 5% of patients who develop STDR within a year would be allocated to a 2-year screening interval. A key limitation of such a rule becomes evident if the reduction in the number of eye examinations (compared to annual screening) is not sufficient to cope with the steady increase in the number of PWD given the lack of extra funding for screening. In particular, there are a number of limitations associated with this rule: (i) our data suggest that this rule lacks of accuracy to identify low-risk patients (56% specificity, Figure 3, left panel) and 44% of patients who do not develop STDR would be invited to annual screening, which affects a considerable number of patients given the low incidence rate of STDR. Multivariate risk-based models are likely to offer a more cost-effective solution. The statistical risk model in (38) and the one presented here could generate a reduction in the number of screening episodes of above 40%, while achieving *acceptable* values of sensitivity. (ii) *Acceptability to patients and staff*. This is an important aspect that has not yet been evaluated. The acceptability of variable interval screening in a pragmatic whole population based on the impact on attendance rates to screening should be taken into account (this consideration applies to any model or rule considered for implementation). (iii) *Long term implementation*. Since two successive annual screening episodes are required for stratification, patients with one of the screening episodes missing will have to be allocated to the high-risk group making the allocation less efficient.

A simpler stratification rule is the one that considers the level of retinopathy of the screening episode at the time of prediction only. Despite its simplicity, it is effective in reducing the number of screening episodes (showing a reduction of 39% compared to a reduction of 27% with the first rule) at the cost of a drop in sensitivity by 7% (88% sensitivity). The recommendation on screening for referable retinopathy by the Scottish Intercollegiate Network (SIGN, updated version from 2014) is based on this second simple rule (although the first rule is expected to be soon introduced in Scotland).

Figure 3 compares the accuracy of the two-episode stratification rule and our multivariate approach. While the two-episode stratification rule identified more patients who developed STDR within a one-year screen interval when compared to the multivariate risk model (95% vs 85%), it also identified fewer patients who did not develop STDR (56% vs 84%). Given the low annual incidence rate of STDR (about 2.5%), the low specificity of the two-episode stratification rule would lead to a very low positive predictive value (PPV =5%). In other words, the majority of patients classified as developing STDR within a year would not develop STDR (false positives, yellow areas in Figure 3). With the multivariate risk model the PPV doubles.

We jointly modeled clinical data and retinopathy to accurately predict STDR. A substantial body of evidence suggests that changes in the values of certain risk factors has a beneficial effect on outcomes in diabetic retinal diseases (3, 27 and references therein, 42, 43). The multivariate predictive model we have developed uses baseline clinical data to model changes in DR (transitions among the states no DR and mild NPDR / BDR in either eye). We conclude that long-term progression of DR is driven by the patient’s overall clinical profile with respect to diabetes control and that a risk prediction model using systemic risk factor data, as well as retinopathy level, may offer a better trade-off between achieving an acceptable sensitivity while also keeping a desirable specificity.

**Acknowledgements**. M.G.F, D.M.H and S.P.H acknowledge support from the UK Medical Research Council (Research project MR/L010909/1). M.G.F. also acknowledges support of the UK EPSRC grant EP/N014499/1. This manuscript presents independent research funded by the National Institute for Health Research (NIHR; RP-PG-1210-12016). The views expressed are those of the authors, not those of the UK National Health Service, NIHR or Department of Health. The authors are grateful to the Liverpool ISDR Study Group.

**References**

1. Wong TY, Mwamburi M, Klein R, Larsen M, Flynn H, Hernandez-Medina M, et al. Rates of progression in diabetic retinopathy during different time periods: a systematic review and meta-analysis. Diabet Care 2009;32 :2307– 13. <http://dx.doi.org/10.2337/dc09-0615>
2. Scanlon PH. The English national screening programme for sight-threatening diabetic retinopathy. J Med Screen 2008;15:1–4
3. Klein R, Knudtson MD, Lee KE, Gangnon R, Klein BE. The Wisconsin Epidemiologic Study of Diabetic Retinopathy: XXII the twenty-five-year progression of retinopathy in persons with type 1 diabetes. Ophthalmology 2008;115 :1859– 68. <http://dx.doi.org/10.1016/j.ophtha.2008.08.023>
4. Javitt JC, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. Ann Intern Med 1996;124 :164– 9. <http://dx.doi.org/10.7326/0003-4819-124-1_Part_2-199601011-00017>
5. ETDRS. Photocoagulation for diabetic macular edema. Early Treatment Diabetic Retinopathy Study report number 1. Early Treatment Diabetic Retinopathy Study research group. Arch Ophthalmol1985;103 :1796– 806. <http://dx.doi.org/10.1001/archopht.1985.01050120030015>
6. Savolainen EA, Lee QP. Diabetic retinopathy - need and demand for photocoagulation and its cost-effectiveness: evaluation based on services in the United Kingdom. Diabetologia 1982;23 :138– 40. <http://dx.doi.org/10.1007/BF01271176>
7. DRS. Photocoagulation treatment of proliferative diabetic retinopathy. Clinical application of Diabetic Retinopathy Study (DRS) findings, DRS Report Number 8. The Diabetic Retinopathy Study Research Group. Ophthalmology 1981;88 :583– 600.
8. 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(11):2174-2182
9. 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:No.74.
10. Stratton IM, Aldington SJ, Farmer AJ, Scanlon PH. Personalised risk estimation for progression to sight-threatening diabetic retinopathy: how much does clinical information add to screening data? Diabet Med 2014;31(Suppl1):23–24.
11. Yau JWY, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care 2012;35:556–564,
12. Gallego PH, Craig ME, Hing S, Donaghue KC. Role of blood pressure in development of early retinopathy in adolescents with type 1 diabetes: A prospective cohort study. *BMJ* 2018; 37, a918.
13. 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
14. Younis N, Broadbent DM, Vora JP, Harding SP; Liverpool Diabetic Eye Study. Incidence of sight-threatening retinopa- thy in patients with type 2 diabetes in the Liverpool Diabetic Eye Study: a cohort study. Lancet 2003;361:195–200
15. Stratton IM, Kohner EM, Aldington SJ, Turner RC, Holman RR, Manley SE, et al. UKPDS 50: risk factors for incidence and progression of retinopathy in Type II diabetes over 6 years from diagnosis. Diabetologia 2001;44:156–63. http://dx.doi.org/10.1007/s001250051594
16. Davis MD, Fisher MR, Gangnon RE, et al. Risk factors for high-risk proliferative diabetic retinopathy and severe visual loss: ETDRS Report #18. *Invest Ophthalmol Vis Sci*, 1998;39, 232-252.
17. 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.
18. Agardh CD, Agardh E, Torffvit O. The association between retinopathy, nephropathy, cardiovascular disease and long-term metabolic control in type 1 diabetes mellitus: A five year follow-up study of 442 adult patients in routine care. Diabetes Res Clin Pract, 1997; 35, 113-121.
19. 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. *New England Journal of Medicine*, 1993;329, 977-986.
20. 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;100,1133-39.
21. Klein R, Klein BE, Moss SE, et al. Glycosolated hemoglobin predicts the incidence and progression of diabetic retinopathy. *JAMA* 1988;260, 2864-2871
22. Diabetes Control and Complications Trial Research Group. The effect of intensive diabetes treatment on the progression of diabetic retinopathy in insulin-dependent diabetes mellitus. Arch Ophthalmol 1995;113(1):36-51.
23. Miljanovic b, Glynn Rj, Nathan DM, Manson jE, Schaumberg DA. A prospective study of serum lipids and risk of diabetic macular edema in type 1 diabetes. Diabetes Care 2004;53(11):2883-92.
24. Sampson, C. J., James, M., Broadbent, D. M., & Harding, S. P. Stratifying the NHS Diabetic Eye Screening Programme: into the unknown? Diabetic Medicine; 2016;33(12), 1612-1614. doi:10.1111/dme.13192.
25. 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. http://dx.doi.org/10.2337/dc12-0625
26. UK National Screening Committee Screening for Diabetic Retinopathy. Extending diabetic eye screening intervals for people at low risk of developing sight threatening retinopathy. <https://legacyscreening.phe.org.uk/policydb_download.php?doc=546>. Last accessed 26 January 2018.
27. Scottish Intercollegiate Guidelines Network (SIGN). Management of Diabetes. A National Clinical Guideline . Edinburgh: Royal College of Physicians; 2010. URL: <http://www.sign.ac.uk/assets/sign116.pdf>. Last accessed 27 January 2018.
28. Public Health England. NHS public health functions agreement 2017-2018. Service specification no.22 NHS Diabetic Eye Screening Programme. <https://www.england.nhs.uk/wp-content/uploads/2017/04/service-spec-22.pdf>. Last accessed 22 February 2018.
29. Hughes D, Komárek A, Bonnett L, Czanner G, García-Fiñana M. Dynamic longitudinal discriminant analysis using multiple longitudinal markers of different types. Statistical Methods in Medical Research 2018; 27(7):2060-2080
30. Hughes D, Komárek A, Czanner G, García-Fiñana M. Dynamic prediction using credible intervals in longitudinal discriminant analysis. Statistics in Medicine 2017; 36:3858-3874
31. Plummer M. Penalized loss functions for Bayesian model comparison. Biostatistics 2008;9:523-539
32. Komárek A, Komárková L. Capabilities of R Package mixAK for Clustering Based on Multivariate Continuous and Discrete Longitudinal Data. Journal of Statistical Software 2014:59(12), 1-38.
33. American Diabetes Association. American Diabetes Association: Standards of Medical Care in Diabetes. Diabet Care 2013; 36:S11– 66. <http://dx.doi.org/10.2337/dc13-S011>
34. Olafsdottir E, Stefansson E. Biennial eye screening in patients with diabetes without retinopathy:10-year experience. Br. J. Ophthalmol 2007;91:1599–601.

<http://dx.doi.org/10.1136/bjo.2007.123810>

1. Aspelund T, Thornorisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia. 2011;54(10):2525–2532. doi: 10.1007/s00125-011-2257-7. [[PubMed](https://www.ncbi.nlm.nih.gov/pubmed/21792613)] [[Cross Ref](https://dx.doi.org/10.1007%2Fs00125-011-2257-7)]
2. Scanlon PH. Screening intervals for diabetic retinopathy and implications for care. Curr Diab Rep. 2017;17(10): 96
3. Nathan DM, Bebu I, Hainsworth D, Klein R, Tamboriane W, Lorenzi G, Gubitosi-Klug R, Lachin JM (DCC/EDIC Research group). Frequency of evidence-based screening for retinopathy in Type 1 Diabetes. New England Journal of Medicine 2017;376(16):1507-1516.
4. Aspelund T, Þórisdóttir O, Ólafsdottir E, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia 2011;54:2525-2532
5. van der Heijden AA, Walraven I, van’ t Riet E, Aspelund T, Lund SH, Elders P, et al. Validation of a model to estimate personalised screening frequency to monitor diabetic retinopathy. Diabetologia 2014;57:1332– 8. http://dx.doi.org/10.1007/s00125-014-3246-4
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;56:1716–1725.
7. Cichosz SL, Johansen MD, Knudsen ST, Hansen TK, Hejlesen O. A classification model for predicting eye disease in newly diagnosed people with type 2 diabetes. Res Clin Pract. 2015;108(2):210-215.
8. Mehlsen J, Erlandsen M, Poulsen PL, Bek T. Individualized optimization of the screening interval for diabetic retinopathy: a new model. Acta Ophthalmol 2012;90 :109– 14. <http://dx.doi.org/> 10.1111/j.1755-3768.2010.01882.x
9. Fong DS, Aiello LP, Ferris FL III, Klein R. Diabetic retinopathy. Diabet Care 2004;27:2540–53. <http://dx.doi.org/10.2337/diacare.27.10.2540>

**Legends to figures**

**Figure 1**. Boxplots for the predicted probability of developing STDR within a year when applying our multivariate discriminant model (left panel), ROC curve (right panel) with sensitivity (85.4%) and specificity (84.0%) illustrated by the green dot.

**Figure 2**. Boxplots for the predicted probability of developing STDR within a year for patients with no DR at the time of prediction (left), mild non-proliferative/background DR (mild NPDR / BDR) in one eye (middle), and mild NPDR / BDR in both eyes (right). The boxplots display the distribution of the data (minimum, first quartile, median, third quartile, and maximum). Values greater than 1.5 times the third quartile are shown separately as plotted points (outliers).

**Figure 3**. Prediction accuracy of the two-episode stratification rule versus our personalized

risk model.

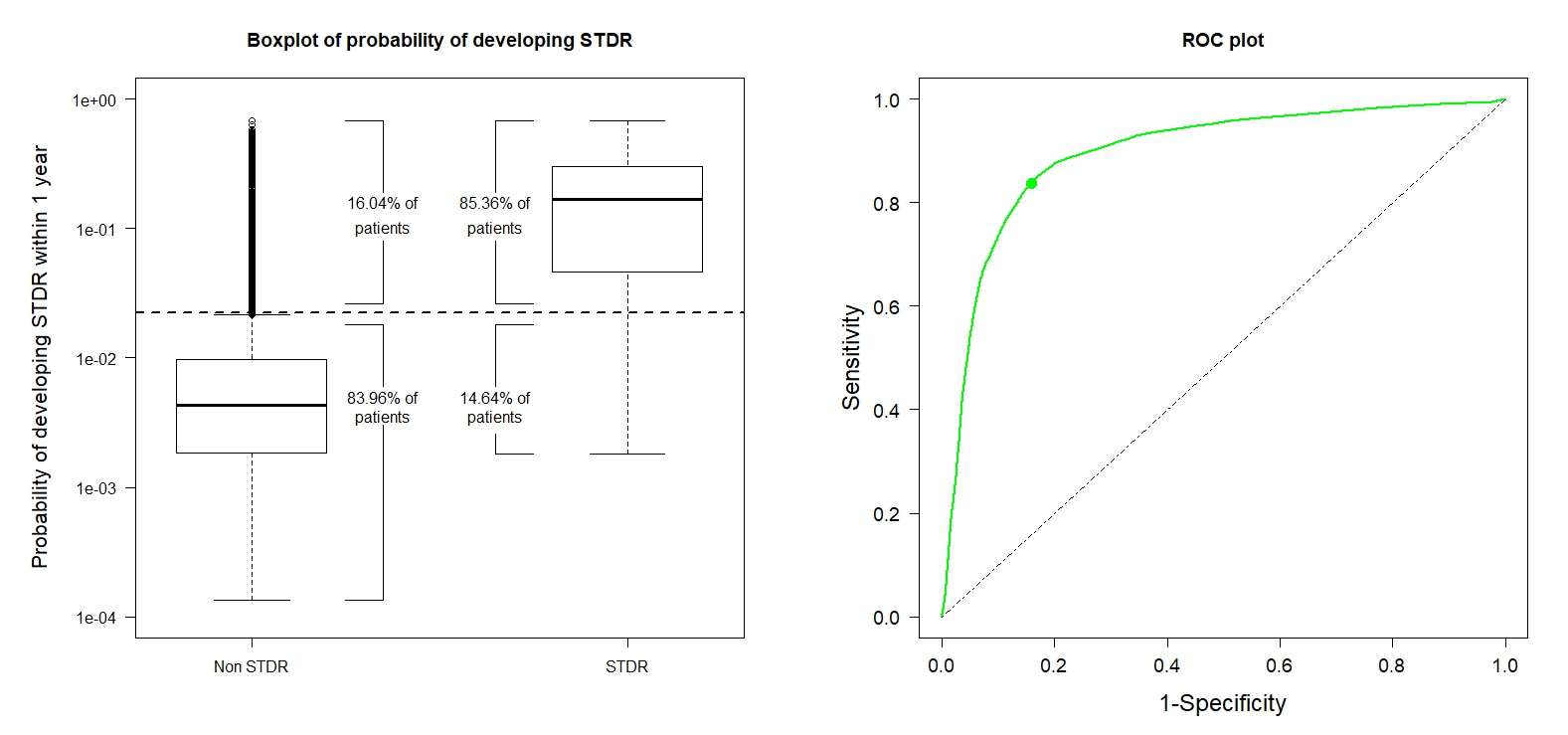
**Tables**

Table 1. Summary measures of demographic, clinical and level of retinopathy data. Mean (standard deviation) (\*) and median (inter-quartile range) (\*\*) are reported as measures of location and variability. Retinopathy grades R0=non-DR and R1=mild non-proliferative/background DR.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **All Patients** | **Non-STDR group** | **STDR group** |
| Number of patients | 13,103 | 12,762 | 341 |
| Total screening visits | 49,520 | 48,562 | 958 |
| Sex (Female) (%) | 5469 (41.7%) | 5350 (41.9%) | 119 (34.9%) |
| Age at first visit\*, years | 59.47 (13.34) | 59.65 (13.27) | 52.93 (14.43) |
| Type I diabetes (%) | 651 (5.0%) | 585 (4.6%) | 66 (19.4%) |
| Duration of diabetes\*\*, years | 1.95 (0.29,4.73) | 1.84 (0.28,4.59) | 4.95 (2.26,9.28) |
| Follow-up length\*\*, years | 6.18 (3.33,8.84) | 6.20 (3.32,8.89) | 5.51 (3.66,7.30) |
| Ethnicity (white:non white:not reported) | 76%:6%:18% | 76%:6%:18% | 69%:13%:18% |
| HbA1c\*\* (mmol/mol)  HbA1c\*\* (%) | 51 (44,60) | 50 (44,59) | 66 (53,87) |
| 6.8% (6.2%,7.6%) | 6.7% (6.2%,7.5%) | 8.2% (7%,10.1%) |
| Cholesterol\* (mmol/L) | 4.18 (1.01) | 4.18 (1.01) | 4.25 (1.07) |
| DBP value\* (mm/Hg) | 75.2 (9.15) | 75.16 (9.15) | 77.09 (9.18) |
| SBP value\* (mm/Hg) | 131.52 (14.08) | 131.47 (14.06) | 134.27 (15.04) |
| HDL value\* (mmol/L) | 1.26 (0.37) | 1.26 (0.37) | 1.23 (0.41) |
| LDL value\* (mmol/L) | 2.1 (0.85) | 2.1 (0.85) | 2.2 (0.84) |
| eGFR value \*\*(mL/min/1.73m2) | 76 (63,88) | 76 (63,88) | 84 (72,90) |
| Missed appointment at previous visit before prediction | 631 (4.8%%) | 551 (4.3%) | 80 (23.5%) |
| Retinopathy grades at first visit (R0/R0:R0/R1:R1/R1) | 75%:16%:9% | 77%:16%:7% | 19%:22%:59% |
| Retinopathy grades at prediction visit (R0/R0:R0/R1:R1/R1) | 78%:13%:9% | 79%:13%:8% | 12%:14%:74% |

**Figures**

**Figure 1**



**Figure 2**



**Figure 3**

