Predictors for diabetic retinopathy progression—findings from nominal group technique and Evidence review

Sajjad Haider 1,2,3, Salman Naveed Sadiq,2 Eniya Lufumpa,1 Harpreet Sihre,1 Mohammad Tallouzi,1 David J Moore,3 Krishnarajah Nirantharakumar,3 Malcolm James Price3

ABSTRACT

Objectives Risk stratification is needed for patients referred to hospital eye services by Diabetic Eye Screening Programme UK. This requires a set of candidate predictors. The literature contains a large number of predictors. The objective of this research was to arrive at a small set of clinically important predictors for the outcome of the progression of diabetic retinopathy (DR). They need to be evidence based and readily available during the clinical consultation.

Methods and analysis Initial list of predictors was obtained from a systematic review of prediction models. We sought the clinical expert opinion using a formal qualitative study design. A series of nominal group technique meetings to shorten the list and to rank the predictors for importance by voting were held with National Health Service hospital-based clinicians involved in caring for patients with DR in the UK. We then evaluated the evidence base for the selected predictors by critically appraising the evidence.

Results The source list was presented at nominal group meetings (n=4), attended by 44 clinicians. Twenty-five predictors from the original list were ranked as important predictors and eight new predictors were proposed. Two additional predictors were retained after evidence check. Of these 35, 21 had robust supporting evidence in the literature condensed into a set of 19 predictors by categorising DR.

Conclusion We identified a set of 19 clinically meaningful predictors of DR progression that can help stratify higher-risk patients referred to hospital eye services and should be considered in the development of an individual risk stratification model.

Study design A qualitative study and evidence review.

Setting Secondary eye care centres in North East, Midlands and South of England.

INTRODUCTION

Diabetic retinopathy (DR) can develop in anyone with diabetes mellitus (DM) and is a major cause of blindness due to damage to, and disruption of the retina (the light-sensitive layer at the back of the eye) leading to loss of sight. DR is a consequence of changes to the blood vessels in this part of the eye leading to leakage of blood and fluid and the formation of abnormal blood vessels.1

There has been a global increase in the number of adults with DM from an estimated 108million in 1980 to 422million in 2014.2 The rise in diabetes prevalence coupled with early detection of DR through better population eye screening coverage has increased the burden of patients with DR to healthcare systems.3

In the UK, DR services are organised into the Diabetic Eye Screening Programme (DESP) for lower-risk patients and the hospital eye services (HES) for higher-risk patients with referable DR, with the HES providing treatment, and closer observation of patients. While services may be organised differently in other countries, the care pathways are likely to be similar for these patients. Patients with DR are referred to the HES when they develop clinical signs of the so called sight-threatening retinopathy. The clinical signs based on photographs are the only differentiating features used for this risk stratification. However, approximately 50%–70% of referrals5 do not need intervention and are observed in
the HES for varying periods of time. This is one reason for a demand and capacity mismatch in HES.

There have been successful attempts to optimise the diabetic screening services through risk stratification of patients with the help of a prognostic prediction model. A similar approach could be used to improve safety and efficiency of DR services in the hospital environment. However, a recent systematic review of prediction models for DR found that none of the 14 models identified were directly applicable for the high-risk patients in the hospital setting.

Predictors are at the core of prevention and prediction of a clinical event. Predictors could be, for example, an individual attribute, a clinical feature, a physiological, psychosocial or an environmental factor. Predictor research ranges from ‘exploration’ to ‘confirmation’ to ‘replication’. Exploration refers to a predictor being mentioned in a primary study as a risk factor for being part of the causal pathway. Confirmation is established if a predictor retains prognostic value even after adjustment for other predictors. Replication is the assessment of the predictor in multiple independent studies. Publication and reporting biases are common in predictors’ research.

A prediction model combines two or more predictors (also called prognostic factors) to predict the likelihood of an outcome, for example, DR progression to a treatment requiring stage or loss of vision, before it occurs. Previous prediction models in the systematic review have between them used 78 different candidate predictors. However, there are problems with the direct use of these predictors for the HES setting. First, a set of predictors this large is not feasible for use in clinical practice. Second, many of these predictors have not been confirmed. Third, a number of predictors were extrapolated from evidence for macrovascular outcomes like stroke to retinopathy progression inappropriately, for example, ‘smoking’ and ‘ethnicity’. Fourth, there was significant duplication and overlap between the predictors, for example, diabetic nephropathy and chronic kidney disease. Such highly correlated predictors are not be independently predictive in the same model. Finally, there may be predictors for high-risk patients, not reported in the prediction model studies as they were primarily for low-risk patients.

There were two comprehensive and up to date evidence reviews on the subject of predictors. However, their perspective was quite different from ours. They were a rich resource of evidence and very useful for horizon scanning, but most of the predictors they suggested are still in research domain and not being recorded in the patient notes, so not possible

**Figure 1** List of candidate predictors from systematic review model development studies. Pie chart illustrates the percentage proportion of each category of predictors. Boxes indicate the type of candidate predictors in each category along with the number of modelling studies that considered them. Information on the full list of candidate predictors is given in online supplemental appendix A3. The predictor was used in at least one model development. eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; HDL, High Density Lipoproteins; LDL, Low Density Lipoproteins; WBC, White Blood Cells; BDR, Background Diabetic Retinopathy; NPDR, Non Proliferative Diabetic Retinopathy; PDR, Proliferative Diabetic Retinopathy.

Aims
In this qualitative study, we aimed to deduplicate the list of predictors and through clinicians’ input and evidence assessment to arrive at a shorter list, more clinically relevant to high-risk patients.

Objectives
1. To identify all the predictors for DR progression in the literature.
2. To seek clinicians’ opinion about the most important predictors among them for high-risk patients and to identify any new predictors not yet described.
3. To assess the evidence base from published literature for each predictor identified in objective 2.
4. To finalise a potentially parsimonious set, that is, a minimal number of predictors in a future multivariate model able to give the highest predictive performance.14 15

METHODS
Study design and methods
Our recent systematic review of prognostic modelling studies for the development of DR of any type including maculopathy in patients with diabetes7 was the primary source of a list of candidate predictors. The text and the reference lists of the modelling studies were searched for the candidate predictors. A qualitative study design (Delphi) has been used with the aim to prioritise the list of already identified predictors.16

Techniques other than nominal group technique (NGT) such as brainstorming, focus group and Delphi methods were considered17 along with their technical strengths and

Table 1  List of predictors chosen by NGT by number of votes and voting frequency

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Voting frequency</th>
<th>Proportion voted (total n=44) (%)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 HbA1c</td>
<td>31</td>
<td>70</td>
</tr>
<tr>
<td>2 Duration of diabetes</td>
<td>24</td>
<td>55</td>
</tr>
<tr>
<td>3 Retinopathy level</td>
<td>17</td>
<td>39</td>
</tr>
<tr>
<td>4 Townsend score</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td>5 Smoking (lifestyle)</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>6 Race</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>7 Proliferative DR</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>8 DNA†</td>
<td>10</td>
<td>23</td>
</tr>
<tr>
<td>9 Nephropathy</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>10 Hypertension</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>11 Maculopathy</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>12 Pregnancy†</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>13 Comorbidities†</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>14 Exercise/physical activity</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>15 Type of DM</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>16 BMI</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>17 eGFR</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>18 Chronic renal disease</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>19 Rapid reduction of blood sugar†</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>20 Dyslipidaemia</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>21 Psychiatric illness</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>22 Visual acuity score</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>23 DBP</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>24 Only eye†</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>25 Age at diagnosis</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>26 Chronic infection†</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>27 Preproliferative†</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>28 Neuropathy</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>29 Age</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>30 Statins</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>31 Insulin</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>32 Gender</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>33 Diet†</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

*Voting frequency in percentages arranged in order of high to low. NGT selected 25 out of 78 candidate predictors list. †Eight new predictors were added by NGT for a total of 33 predictors.

BMI, body mass index; DBP, diastolic blood pressure; DM, diabetes mellitus; DNA, did not attend; DR, diabetic retinopathy; eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; NGT, nominal group technique.
limitations. NGT was used to shrink the list of candidate predictors. Primary studies were then evaluated for evidence. In case no reference of the primary study was found in the modelling study, searches were made in Pubmed and Scopus. Subsequently, the evidence behind the shortlisted predictors was evaluated using the Quality in Prognosis Studies (QUIP) tool. To comply with the PROGRESS (Prognosis research strategy) framework for predictors, we used a criterion of at least ‘exploration’ (exploring predictor’s relation to prognosis) within primary studies before including them into the list of our candidate predictors for future models, to ensure that the risk of the outcomes of interest can be calculated more precisely by using those predictors by a future model.

Nominal group technique
NGT is a qualitative research methodology, where every meeting is a structured small group exercise allowing judgments by individuals to be pooled to arrive at a decision in an uncertain situation. NGT has been used frequently in ophthalmology and medicine and is a highly adaptable method.

An information pack with the predictors list from the systematic review was sent to participating unit before the meeting (online supplemental appendices A1–A4). Informed consent forms were provided and signed by participants (online supplemental appendix A2). NGT was performed through a series of four meetings, each lasting around 1.5 hours. Meetings were accommodated within clinical governance/audit meetings of four National Health Service (NHS) trusts with the permission of respective research and development directorates. NGT was chosen, because of its modifiable nature allowing brainstorming to decide on importance of predictors, a round-robin for an equal opportunity for all participants, a discussion for clarifications, voting for final decision on ranking and for being time efficient. Each meeting was conducted in four stages (online supplemental appendix A4): from providing background information, round-robin recording of ideas, discussion of the list of ideas and ending in voting for ranking of the ideas generated.

During the round-robin recording, the facilitator went around the table inviting one item from each participant at a time (to give equal opportunity to all the participants) and recording them on the flip chart. It was left open to participants to choose any number of predictors from the list provided or name their own predictors using their personal experience and insight. We requested participants to restrict their final choices to a total of 15–20 predictors.

The round-robin cycles (within a meeting and between the meetings) were repeated until saturation had been achieved—a point where no new predictors were being added and thus all new information had been obtained. The centres were recruited sequentially and from the second meeting onward, after every meeting cumulative results were examined for any new predictors suggested and thus monitoring for saturation (when no further predictors were being added) before stopping to recruit.

Participants
To ensure a maximum range of views and opinions were collected, keeping purposeful sampling in mind that is, selection of information-rich resource for the research question as a priority, we approached medical retina team leaders for four NHS trusts (consultants with much longer training and experience) with their teams of middle grades and registrars, nurses and optometrists, directly involved in caring for patients with DR, for these meetings. We aimed to over-recruit allowing for likely no-shows/dropouts on the day for each group.

Data collection
Reflexivity of the authors was considered when designing the project. SH is an ophthalmologist and thus shared the participants’ experience. He moderated the meetings, but one or more of the qualitative study researchers (HS, EL or MT) were also present to reduce the chances of bias. Clinicians were asked for their written consent (online supplemental appendix A2) for information collected to be used for further research and publication. Flip chart and marked lists were collected at the end of discussion while voting and all other information was collated on a spreadsheet. All participant data were anonymised, with no direct quotes mentioned in any dissemination.

Data extraction in spreadsheet and collation after every meeting were carried out by one of the coauthors involved in that meeting and validated with the paper forms with the help of a second researcher. We collated all new predictors suggested by clinical colleagues participating in these meetings, added them to the list provided and helped participants rank them for importance in prognosis prediction.

Analyses of NGT
Voting frequency was calculated for voting scores per predictor (summing of scores) as well as frequency of voting percentage (score achieved for the item/maximum possible score × 100) and tabulated according to their rank.

Evidence review
In this step, we searched the reference lists of the modelling studies identified in the systematic review for primary studies that had investigated the predictors selected by NGT and subjected them to critical appraisal using the QUIP tool. A basic literature search was then performed in Pubmed and Scopus to identify primary studies for predictors where no primary evidence source was referenced in any modelling study. Using the identified information, we summarised the status of each predictor as explored, confirmed or confirmed and replicated in more than one study.

Patients and public involvement
While there was no direct patients and public involvement required in this research, we included all the predictors from Individualised Screening for Diabetic Retinopathy
Evidence given a total of 35 predictors (table 2). ‘Systolic blood pressure’ (SBP) was included because of the participation of patient expert panel in ISDR model. The NGT participants implicitly included it by selecting hypertension as a predictor for DR and is more sensitive than the earlier ‘estimated glomerular filtration rate’ (eGFR). Therefore eGFR was retained in preference to ‘chronic kidney disease’ and ‘diabetic nephropathy’ as it is established as a predictor for DR and is more sensitive than the earlier two predictors. In the case of the overlap between ‘hypertension’, ‘DBP’ and ‘SBP’, the latter two predictors are represented in ‘hypertension’ but seem to have different prognostic values, so were retained.

We did not find any primary studies supporting the association between DR progression and ‘psychiatric illness’ or ‘diabetic neuropathy’. Among the studies found, no proven association of DR with ‘BMI’, ‘exercise/physical activity’ or ‘smoking’ was seen but a weak association was seen with ‘serum cholesterol’. The last item was therefore retained, and the rest excluded. Among the new predictors suggested by NGT, primary studies were found confirming the association for ‘early worsening’, ‘pregnancy’, frequent ‘DNA’, and ‘preproliferative DR’. Due to the lack of evidence, the remaining four predictors (‘comorbidities’, ‘only eye situation’, ‘Diet’ and ‘chronic infection’) were excluded.

The most common primary studies quoted were UKPDS, Diabetes Control and Complications Trial (DCCT), Wisconsin Epidemiological Study of Diabetic Retinopathy, Epidemiology and Prevention of Diabetes Study (EURODIAB) and Action to control cardiovascular risk in diabetes (ACCORD). Table 2 tabulates primary studies and status of confirmation and replication for individual predictors. Fourteen predictors were excluded for reasons of no proven association, duplication/overlap. The remaining 21 predictors were condensed into 19 predictors (table 3) by combining the DR categories together.

Only 2 out of 14 modelling studies reported primary studies for all of the prognostic factors/predictors used. One of the modelling studies used the items without giving reference to any primary study, but instead quoted literature review/expert patients’ panel. A model partly relied on borrowing the items from other models, some could be traced from the reference list but were not in context.

Most of the primary studies had multiple publications. Out of 35 predictors, 25 (71%) had good supporting evidence of predictive value from the literature (table 2). Four predictors were excluded because of overlap/duplication. Out of 18 primary studies critically appraised in online supplemental appendix A7, the evidence base for only one primary study had a high risk of bias, mainly due to confounders issues and highly selective population. The factor involved did have another supporting study with low risk of bias. Three predictor studies were judged to have moderate risk of bias and 15 predictor studies were low risk of bias (79% of primary predictors studies) on QUIP criteria.

DISCUSSION

Statement of principal findings

We used clinical opinion in the NGT meetings to reduce a list of 78 previous candidate predictors to a list of 25 potential predictors. The study also suggested a further

Ethics

The protocol was evaluated by the four NHS trusts where the NGT was carried out. Since the research was essentially a decision exercise based on clinicians’ expert opinion and did not involve patient data, no ethical approval was required.

RESULTS

The list of 78 candidate predictors from the systematic review are given in figure 1 (shows their breakdown and frequency of their use in the pool of 14 models) and in Appendix A3. Biochemical predictors were used most commonly. Figure 2 shows the predictor items flow during the processes of NGT and evidence evaluation.

Nominal group technique

We conducted four NGT meetings, within secondary care eye centres in NHS trusts in the following areas in the UK: (1, 2) Two Midland trusts, (3) North East England and (4) South of England. After the fourth NGT meeting, saturation was reached. The full set of 33 predictors selected by the clinicians are shown in table 1. Participants’ roles varied from consultants, specialist registrars, middle-grade doctors and allied health professionals (one optometrist and one nurse). A total of 44 clinicians participated with numbers per session ranging from 6 to 16 (online supplemental appendix A4). Eleven out of 19 participating consultants were medical retina trained. After multiple round-robin cycles (within a meeting and different meetings) had been repeated to ensure saturation had been achieved and no new predictors were being added (online supplemental appendix A6), the final set of candidate predictors as decided by NGT is given in table 1.

Evidence evaluation

Following NGT, during the evidence review, two further predictors from the original list of 78 were added back to give a total of 35 predictors (table 2). ‘Systolic blood pressure’ (SBP) was included because of the participation of patient expert panel in ISDR model, was the third the most common predictor used in prognostic models and has a good evidence base. The NGT participants probably implicitly included it by selecting hypertension as well. Total serum cholesterol was added back as a proxy for dyslipidaemia as difficulties with reporting of this predictor have previously been reported.

After NGT, residual duplication/overlap still remained, for example, ‘diabetic nephropathy’ and ‘chronic kidney disease’, ‘hypertension’, ‘diastolic blood pressure’ (DBP) and ‘SBP’, ‘dyslipidaemia’, ‘cholesterol’ and ‘statins’. ‘Diabetic nephropathy’ is the primary cause of ‘chronic kidney disease’ characterised by progressive decline of kidney disease’ and ‘diabetic nephropathy’ as it is established as a predictor for DR and is more sensitive than the earlier two predictors. In the case of the overlap between ‘hypertension’, ‘DBP’ and ‘SBP’, the latter two predictors are represented in ‘hypertension’ but seem to have different prognostic values, so were retained.

We did not find any primary studies supporting the association between DR progression and ‘psychiatric illness’ or ‘diabetic neuropathy’. Among the studies found, no proven association of DR with ‘BMI’, ‘exercise/physical activity’ or ‘smoking’ was seen but a weak association was seen with ‘serum cholesterol’. The last item was therefore retained, and the rest excluded. Among the new predictors suggested by NGT, primary studies were found confirming the association for ‘early worsening’, ‘pregnancy’, frequent ‘DNA’, and ‘preproliferative DR’. Due to the lack of evidence, the remaining four predictors (‘comorbidities’, ‘only eye situation’, ‘Diet’ and ‘chronic infection’) were excluded. The most common primary studies quoted were UKPDS, Diabetes Control and Complications Trial (DCCT), Wisconsin Epidemiological Study of Diabetic Retinopathy, Epidemiology and Prevention of Diabetes Study (EURODIAB) and Action to control cardiovascular risk in diabetes (ACCORD). Table 2 tabulates primary studies and status of confirmation and replication for individual predictors. Fourteen predictors were excluded for reasons of no proven association, duplication/overlap. The remaining 21 predictors were condensed into 19 predictors (table 3) by combining the DR categories together.

Only 2 out of 14 modelling studies reported primary studies for all of the prognostic factors/predictors used. One of the modelling studies used the items without giving reference to any primary study, but instead quoted literature review/expert patients’ panel. A model partly relied on borrowing the items from other models, some could be traced from the reference list but were not in context. It was not clear in 8 (23%) out of 35 predictors, as to which primary studies confirmed them (given in table 2).

Most of the primary studies had multiple publications. Out of 35 predictors, 25 (71%) had good supporting evidence of predictive value from the literature (table 2). Four predictors were excluded because of overlap/duplication. Out of 18 primary studies critically appraised in online supplemental appendix A7, the evidence base for only one primary study had a high risk of bias, mainly due to confounders issues and highly selective population. The factor involved did have another supporting study with low risk of bias. Three predictor studies were judged to have moderate risk of bias and 15 predictor studies were low risk of bias (79% of primary predictors studies) on QUIP criteria.

DISCUSSION

Statement of principal findings

We used clinical opinion in the NGT meetings to reduce a list of 78 previous candidate predictors to a list of 25 potential predictors. The study also suggested a further
After evidence review, we added back another two predictors. The new predictors suggested by the NGT made good clinical/biological sense, but four of them (‘Comorbidities’, ‘only eye situation’, ‘diet’, ‘chronic infections’) have not been explored for association with DR progression, although it is possible that the last two may be operational through other variables such as glycated haemoglobin (HbA1c) rise due to uncontrolled blood sugar. The ‘only eye’ situation clearly has a higher risk of blindness because of the absence of function in the affected eye, or higher risk because of the same pathology as the

**Table 2** Predictors and their primary studies with confirmation and replication status

<table>
<thead>
<tr>
<th>Group</th>
<th>No</th>
<th>Predictors</th>
<th>Model mentioning predictor</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ocular features</strong></td>
<td>1</td>
<td>Retinopathy level/DR grade</td>
<td>Lagani et al.(^{14})</td>
<td>C, (^{49}) R(^{70})</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Proliferative DR</td>
<td>Hippisley-Cox and Coupland(^{11*})</td>
<td>C, (^{30}) R(^{71})</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Maculopathy</td>
<td>Hippisley-Cox et al.(^{11*})</td>
<td>C, (^{48}) R(^{71})</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Visual acuity score</td>
<td>Lagani et al.(^{14})</td>
<td>C, (^{51})</td>
</tr>
<tr>
<td><strong>Sociodemographics</strong></td>
<td>5</td>
<td>Age</td>
<td>Multiple studies(^{14, 42, 57})</td>
<td>C, (^{43}) R(^{65})</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Race</td>
<td>Harris et al.(^{11, 63}) (L)</td>
<td>E(^{4}) Lack of evidence</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Gender</td>
<td>Harris et al.(^{63})</td>
<td>C, (^{62}) R(^{43})</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Social deprivation score</td>
<td>Hippisley-Cox and Coupland(^{11*})</td>
<td>C, (^{52}) R(^{72})</td>
</tr>
<tr>
<td><strong>Diabetes characteristics</strong></td>
<td>9</td>
<td>Type of DM</td>
<td>Icelandic model(^{73*})</td>
<td>C(^{74})</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Age at diagnosis</td>
<td>UKPDS OM(^{27*})</td>
<td>C, (^{49}) R(^{55})</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Duration of DM</td>
<td>Icelandic model(^{73})</td>
<td>C, (^{48}) R(^{70})</td>
</tr>
<tr>
<td><strong>Biochemical parameters</strong></td>
<td>12</td>
<td>HbA1c</td>
<td>UKPDS OM(^{17*})</td>
<td>C, (^{51}) R(^{49})</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>eGFR</td>
<td>UKPDS OM(^{27*})</td>
<td>C, (^{78}) R(^{38})</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Total serum cholesterol</td>
<td>Soedamah-Muthu et al.(^{67})</td>
<td>C, (^{39}) R(^{55})</td>
</tr>
<tr>
<td><strong>Physical examination</strong></td>
<td>15</td>
<td>DBP</td>
<td>Icelandic model(^{75*})</td>
<td>C(^{76})</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>SBP</td>
<td>UKPDS OM(^{27*})</td>
<td>C, (^{76}) R(^{55})</td>
</tr>
<tr>
<td><strong>Comorbidities</strong></td>
<td>17</td>
<td>Hypertension</td>
<td>Harris et al.(^{63})</td>
<td>C(^{39})</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>Dyslipidaemia</td>
<td>Harris et al.(^{63})</td>
<td>C, (^{73}) R(^{80})</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>Psychiatric illness</td>
<td>Lagani et al.(^{12}) (L)*</td>
<td>Absence of evidence</td>
</tr>
<tr>
<td><strong>Diabetic complications</strong></td>
<td>20</td>
<td>Chronic renal disease</td>
<td>*Hippisley-Cox and Coupland(^{11})</td>
<td>C(^{81})</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>Diabetic nephropathy</td>
<td>Harris et al.(^{63})</td>
<td>C, (^{62}) R(^{37})</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>Diabetic neuropathy</td>
<td>Harris et al.(^{63}) (L)*</td>
<td>Absence of evidence</td>
</tr>
<tr>
<td><strong>Diabetes treatment</strong></td>
<td>23</td>
<td>Statin</td>
<td>Harris et al.(^{63})</td>
<td>C, (^{85}) R(^{80})</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>Insulin</td>
<td>Harris et al.(^{63*})</td>
<td>C, (^{83}) R(^{50})</td>
</tr>
<tr>
<td><strong>Lifestyle</strong></td>
<td>25</td>
<td>Smoking</td>
<td>McEwan et al.(^{61}) (L)</td>
<td>C, (^{44}) (^{34}) Absence of evidence</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>BMI</td>
<td>McEwan et al and others(^{14, 42, 61}) (L)*</td>
<td>C, (^{41}) Absence of evidence</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>Exercise/physical activity</td>
<td>Tanaka(^{13}) (L)</td>
<td>Absence of evidence</td>
</tr>
<tr>
<td><strong>New from NGT</strong></td>
<td>28</td>
<td>Only eye situation</td>
<td>NA (L)</td>
<td>Absence of evidence</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>Early worsening</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>Frequent DNA</td>
<td>NA (L)</td>
<td>C, (^{49}) R(^{84})</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>Pregnancy</td>
<td>NA</td>
<td>C, (^{47}) Lack of evidence</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>Diet</td>
<td>NA (L)</td>
<td>C, (^{49}) R(^{62})</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>Preproliferative DR</td>
<td>NA</td>
<td>C, (^{48}) R(^{85})</td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>Chronic infection</td>
<td>NA (L)</td>
<td>Absence of evidence</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>Comorbidities</td>
<td>NA (L)</td>
<td>Absence of evidence</td>
</tr>
</tbody>
</table>

*Modelling study did not clearly identify the primary study for the predictor. Details in text.

BMI, body mass index; C, confirmation; DBP, diastolic blood pressure; DM, diabetes mellitus; DNA, did not attend; DR, diabetic retinopathy; E, exploration; eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; ISDR, Individualised Screening for Diabetic Retinopathy; L, lacking evidence; NA, not applicable; R, replication; SBP, systolic blood pressure; UKPDS, United Kingdom prospective Diabetes Study.
lost eye but has not been investigated for its association with DR progression. The other three predictors, ‘early worsening’, ‘pregnancy’ and ‘preproliferative DR’ have a proven association with our outcome of interest and were therefore included. While specific comorbidities were considered as mentioned above, ‘comorbidities’ presence or numbers as a predictor had no supporting evidence. ‘Frequent DNA’ has been proven to add prognostic value, so has been added to the final list but is in need for further confirmation.

The top three ranked predictors from NGT (HbA1c, duration of diabetes and DR grade) were the same as the low risk of bias models’ predictor sets (online supplemental appendix A7) chosen by the systematic review.7 This was in spite of this information being withheld from the participants in the NGT meetings. This shows the clinicians’ intuitive thinking matches the findings of the systematic review as well. The approach of using retinopathy stages data alone to develop a risk stratification tool in DESP environment has been suggested.60 ISDR model35 in complete contrast has suggested using clinical predictors alone, especially for higher-risk patients. We looked at risk estimates of the various predictors within various models. They tend to vary depending on the combination of predictors used.6 14 However, ocular predictors generally showed higher relative risk estimates. We suggest that any future model should contain a combination of ocular and systemic predictors for use in higher risk hospital population, as the values of, for example, SBP and HbA1c and so on are likely to be higher and can thus add significantly to the predictive ability. A practical sized set is now available, with individual predictors ranked in importance as perceived by the NGT participants. A prediction model for higher-risk DR could be built based on them in appropriate population data.

Several predictors in previously reported models were associated with diabetic complications other than retinopathy—in multiple outcomes or composite outcome models.14 57 The evidence for association of some of these predictors with DR is unclear. For example, DR was a main covariate for severity of ‘diabetic polyneuropathy’,11 39 but it still remains to be seen if the reverse is also true. ‘Smoking’ as a candidate predictor was included in the majority of the models but made it to the final set only in one of them.51 The primary studies failed to show it as a predictor for DR progression.43 62 Harris et al33 included ethnicity in their final model but HR p value crossed the recommended threshold of >0.2, so was of doubtful statistical significance. It was included in five models as candidate predictor but not made to any of the final set on statistical testing. However, there is indirect evidence of higher risk of south Asians to develop STR and blacks and South Asians among patients under DESP having a higher prevalence of visual impairment.64 We therefore included ethnicity in the final set.

**Strengths and weaknesses of the study**

We based this study on the full list of predictors that had duplication, overlap and extrapolation of evidence from outcomes other than DR progression. In this situation with a large number of predictors already confirmed and replicated, incorporating clinical insight through a qualitative study design was invaluable to reach a shorter, more manageable list and also generate some new predictors. This is also the first study to evaluate the evidence base of potential predictors used in existing prognostic models for DR progression and to follow the Prognosis Research strategy (PROGRESS) framework recommendations.9 We noted reporting deficiencies and have suggested possible preventive strategies for future.

**Strengths and weaknesses in relation to other studies**

The ISDR model35 mentioned using a patient expert panel for decision-making during the predictor selection process. There is not enough detail in the study design, and we assume that it is only comparable to our work in that the study design was a qualitative one. We provide
here the details of our methods and results with the interpretation. Our list also includes all the ISDR predictors, reflecting their expert patients’ input. The Standardisation of Uveitis Nomenclature and Consensus on Outcome Measures for Glaucoma Effectiveness Trials are prime examples of successful use of NGT to arrive at a decision with the help of expert panels in ophthalmology. Another qualitative study looked into the patient-perceived risks of disease and benefits of treatment, but did not address predictors selection. We wanted to build on this approach seeking clinical experts’ opinion with the help of NGT meetings.

Implications for clinicians and policymakers
This set of predictors derived will be useful to risk stratify patients, optimise treatment strategies, inform patients about their personal risks and improve research strategies as well as providing the building blocks for future prognostic models. A set of predictors based on these 19 finalised predictors could be used to risk stratify the population received within the HES after referral from DESP, to help with prioritisation of appointments and thus direct the resources more appropriately. Alternatively, a model could be constructed to estimate an individual patient’s risk. This research will help the clinicians manage their patients according to their risk of progression.

While we have attempted to develop a list of predictors that are useful in predicting patients who progress from referable DR to a stage of needing treatment or vision loss, the list is primarily from patients with diabetes under screening for incident DR and referable DR and as such are generalizable as markers for progression to any stages.

Unanswered questions and future research
Risk of bias assessment did not affect our decision to exclude any predictors as vast majority of the predictors had good evidence base. PROGRESS criteria was used to decide which predictors were not confirmed yet. That has helped decide where there is need for further research, for example, ‘Race’, ‘diet’, ‘exercise’ and so on.

We also found the reporting of the evidence base for the predictors selected in the modelling studies suboptimal. Under-reporting has been mentioned by other observers before and needs improving. The Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement asks for all predictors used to be clearly defined. One way of improving this situation may be for TRIPOD checklist to encourage investigators to report the primary study of origin for every candidate predictor used in a model. Out of 35, 10 (29%) predictors were not supported by the existing evidence base applying PROGRESS standards and require further research.

There have been recent useful reviews and studies on ocular predictors and their natural history use as predictors of DR progression. Ocular signs identified in Optical coherence tomography and Fundus Fluorescein Angiography could also be potential predictors of DR progression. While the ocular signs recommended are suitable for prospective research studies, existing retrospective longitudinal data most commonly used for prognostic research are unlikely to have sufficient information on these predictors. There is an ever-increasing interest in fundus images-based detection, assisted by artificial intelligence. Prognostic factor research is a dynamic field and will benefit greatly with these newer technologies. Machine learning can handle the data from wider sources, can bring additional benefits from automation, unsupervised clustering of a much larger number of predictors and can also add new phenotypes associated with the outcomes. However, there are ethical, governance, interpretability issues and the process of development of these techniques are at an early stage of development and application. It is likely that future research will identify further potentially important predictors so an update may be required in the future.

CONCLUSION
We have been able to identify 19 evidence-based predictors for DR progression, using a novel method (NGT) and evidence review inline with PROGRESS recommendations. This smaller and more practical set provides a useful resource for a potential model to stratify patients for risk of DR progression, to aid clinical decision-making and optimise clinical care pathways. This set is ranked in importance by the NGT.

Acknowledgements The authors would like to thank four NHS trusts and their clinicians for their support and involvement in NGT. Also would like to express gratitude to Sophie Beese for kindly reviewing the manuscript for an internal peer review.

Contributors SH was responsible for concept and design of the work, for acquisition, analysis, interpretation of data and drafting the paper. Rest of the coauthors provided support in above, revised it critically and contributed in final draft.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon request. Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.
REFERENCES


