

The development and validation of a clinical prediction model to determine the probability of MODY in patients with young-onset diabetes

B. M. Shields · T. J. McDonald · S. Ellard ·
M. J. Campbell · C. Hyde · A. T. Hattersley

Received: 16 August 2011 / Accepted: 24 November 2011 / Published online: 5 January 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract

Aims/hypothesis Diagnosing MODY is difficult. To date, selection for molecular genetic testing for MODY has used discrete cut-offs of limited clinical characteristics with varying sensitivity and specificity. We aimed to use multiple, weighted, clinical criteria to determine an individual's probability of having MODY, as a crucial tool for rational genetic testing.

Methods We developed prediction models using logistic regression on data from 1,191 patients with MODY ($n=594$), type 1 diabetes ($n=278$) and type 2 diabetes ($n=319$). Model performance was assessed by receiver operating characteristic

(ROC) curves, cross-validation and validation in a further 350 patients.

Results The models defined an overall probability of MODY using a weighted combination of the most discriminative characteristics. For MODY, compared with type 1 diabetes, these were: lower HbA_{1c}, parent with diabetes, female sex and older age at diagnosis. MODY was discriminated from type 2 diabetes by: lower BMI, younger age at diagnosis, female sex, lower HbA_{1c}, parent with diabetes, and not being treated with oral hypoglycaemic agents or insulin. Both models showed excellent discrimination (c-statistic=0.95 and 0.98, respectively), low rates of cross-validated misclassification (9.2% and 5.3%), and good performance on the external test dataset (c-statistic=0.95 and 0.94). Using the optimal cut-offs, the probability models improved the sensitivity (91% vs 72%) and specificity (94% vs 91%) for identifying MODY compared with standard criteria of diagnosis <25 years and an affected parent. The models are now available online at www.diabetesgenes.org.

Conclusions/interpretation We have developed clinical prediction models that calculate an individual's probability of having MODY. This allows an improved and more rational approach to determine who should have molecular genetic testing.

Keywords Maturity onset diabetes of the young · MODY · Prediction model · Clinical characteristics

Electronic supplementary material The online version of this article (doi:10.1007/s00125-011-2418-8) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

B. M. Shields · A. T. Hattersley (✉)
Peninsula NIHR Clinical Research Facility,
Peninsula Medical School, University of Exeter,
Barrack Road,
Exeter EX2 5DW, UK
e-mail: Andrew.Hattersley@pms.ac.uk

T. J. McDonald · S. Ellard
Royal Devon and Exeter NHS Foundation Trust,
Exeter, UK

M. J. Campbell
School of Health and Related Research, University of Sheffield,
Sheffield, UK

C. Hyde
Peninsula Technology Assessment Group,
Peninsula Medical School, University of Exeter,
Salmon Pool Lane,
Exeter, UK

Abbreviations

GCK Glucokinase
HNF1A Hepatocyte nuclear factor 1- α
HNF4A Hepatocyte nuclear factor 4- α
OHA Oral hypoglycaemic agent
ROC Receiver operating characteristic

Background

MODY is a rare monogenic form of diabetes [1, 2]. Mutations in at least seven genes have been identified to date as causing MODY, with *HNF1A*, *HNF4A* and *GCK* accounting for approximately 94% of cases [3]. A correct diagnosis of MODY has implications for patient treatment: hepatocyte nuclear factor 1- α (HNF1A) and hepatocyte nuclear factor 4- α (HNF4A) MODY are very sensitive to sulfonylurea tablets [4, 5] whereas patients with glucokinase (GCK) MODY rarely require pharmacological treatment [6, 7]. In addition, a diagnosis of MODY predicts disease prognosis and identifies risk of diabetes in family members. It is estimated that more than 80% of individuals with MODY in the UK are currently undiagnosed, or misdiagnosed with type 1 or young-onset type 2 diabetes [3] and, consequently, inappropriately treated, often with insulin.

Genetic testing is expensive, so careful consideration is required when determining which patients should be tested. In young-onset patients, type 1 diabetes is the most prevalent form, accounting for around 74% of cases in those diagnosed under 35 (E. Pearson and R. McAlpine, personal communication). Type 2 diabetes is less common in young patients, but its likelihood increases with age and obesity. MODY is rare, with estimates of its prevalence between 0.3% and 2.4% of diabetes cases [3, 8–11], although it is recognised that these figures represent underestimates.

MODY is characterised by young age at diagnosis, strong family history of diabetes (due to autosomal dominant inheritance), and no insulin dependence [1, 2, 12]. Most clinical criteria for defining MODY have used absolute cut-offs (e.g. non-obese, age at diagnosis <25 years), rather than determining a probability based on continuous quantitative traits such as BMI or age at diagnosis. Criteria based on absolute cut-offs have shown poor sensitivity, picking up approximately only half of patients with MODY [3, 13]. Current guidelines for genetic testing are based on expert knowledge obtained from clinical observation [14], but not all of the characteristics described are routinely collected, and the likelihood of a positive diagnosis based on these clinical features, either separately or in combination, has not been quantified. Identifying a way of combining simple clinical criteria in a weighted manner, to produce a probability of MODY, would be a valuable aid in determining whether to test an individual for MODY.

Medical diagnostic decision models provide a way of determining a person's probability of having a particular condition based on their clinical characteristics. Widely used models such as the Amsterdam criteria and Bethesda guidelines for hereditary non-polyposis colorectal cancer [15] and the Well's score for deep vein thrombosis [16] provide practical screening tools for selecting patients for diagnostic testing. A clinical prediction model that generates a probability of MODY would aid clinicians in identifying patients that are most likely to benefit from genetic testing.

We aimed to produce a clinical prediction model for patients with young-onset diabetes (diagnosed ≤ 35 years) that uses clinical criteria to determine a patient's probability of having MODY compared with the more common type 1 and type 2 diabetes. This can then be turned into a simple web-based algorithm for use by clinicians, patients and the molecular genetics diagnostic laboratory.

Methods

Participants and data

All participants were diagnosed between the ages of 1 and 35. Type 1 diabetes was defined as occurring in patients treated with insulin within 6 months of diagnosis [17]; otherwise patients were defined as having type 2 diabetes. MODY patients were identified based on a confirmed genetic diagnosis of HNF1A, HNF4A or GCK MODY.

Initial dataset

We identified 278 patients with type 1 diabetes criteria and 319 patients with type 2 diabetes criteria from five research databases (see electronic supplementary materials (ESM) Methods 1). These were compared with 594 probands with a genetic diagnosis of MODY (243 GCK, 296 HNF1A and 55 HNF4A) obtained from referrals to the Molecular Genetics Laboratory at the Royal Devon and Exeter NHS Foundation Trust, UK. Of the MODY probands, 177 (30%) were treated with insulin within 6 months of diagnosis and 417 (70%) were not, indicating the proportions of patients who were likely to have been misdiagnosed as having type 1 or type 2 diabetes at referral.

A core dataset comprising key variables common to all databases was established including sex, age at diabetes diagnosis, age at recruitment/referral, BMI, initial and current treatment (diet, oral hypoglycaemic agents [OHAs] or insulin), time to insulin treatment, HbA_{1c} and parent affected with diabetes. As well as including the original reported BMI, for initial exploratory analysis, BMI in children was converted to the equivalent value in adults using the Child Growth Foundation reference standards [18] to enable direct comparison between adults and children. All patients were of white European origin. In the MODY cases, data were taken from the referrals database, thus reflecting the patient characteristics at referral and not following a genetic diagnosis.

Statistical approaches used for the discrimination models

Type 1 and type 2 diabetes are mutually exclusive subtypes according to the definitions used. Therefore, only two-way comparisons were considered: type 1 diabetes vs MODY

and type 2 diabetes vs MODY. Three statistical approaches were used to produce the models: logistic regression, discriminant analysis and classification trees (CART). Further details are provided in ESM Methods 2.

Discriminative ability of the models

From the logistic regression and linear discriminant analysis, fitted probabilities were obtained from the data. Receiver operating characteristic (ROC) curves were plotted to determine the best cut-offs in terms of sensitivity and specificity, and the AUC (c-statistic) was used as a measure of overall performance. For classification trees, model performance was assessed by misclassification rates.

Validation

Jack-knife cross-validation was used on all models as a measure of internal validation. Following completion of the clinical model using the initial dataset, a further 350 patients were identified as a result of ongoing research recruitment/referral: 219 with MODY, of which 83 were probands and 136 were unrelated family members of probands used in the original model, plus a further 90 patients with type 1 diabetes and 41 patients with type 2 diabetes taken from community-based research cohorts (see ESM Methods 1). This served as an external test dataset.

Post-test probabilities of MODY

The pre-test probabilities for MODY were estimated as 0.7% and 4.6% for the type 1 diabetes vs MODY and type 2 diabetes vs MODY models, respectively (see ESM Methods 3 for derivation). Likelihood ratios for MODY were calculated using the sensitivity and specificity obtained from various cut-offs in fitted probabilities in the models. Post-test odds for MODY were obtained by multiplying the pre-test odds by the likelihood ratio, which could then be converted to post-test probabilities (positive predictive values).

Results

Patient characteristics (Fig. 1)

Parental history of diabetes was more common in MODY patients (90% had at least one parent affected vs 61% in type 2 diabetes and 19% in type 1 diabetes) (Fig. 1). The 10% of MODY cases without a parent affected is most likely to reflect cases where the parent has undiagnosed diabetes at referral or, in a minority of cases, where the mutation is de novo. Patients with GCK MODY were more likely to be treated with diet alone (78%), whereas HNF1A/4A MODY

and type 2 diabetic patients were most likely to be treated with OHAs and/or insulin (83% and 91%, respectively). By definition, all patients with type 1 diabetes were treated with insulin. There was a higher proportion of female patients in the MODY groups (65% in GCK MODY, 72% in HNF1A/4A MODY, 53% in type 1 diabetes, 50% in type 2 diabetes).

Figure 1 shows density plots of the continuous characteristics of the four subtypes of diabetes. Type 2 diabetic patients were diagnosed at an older age compared with the other three types of diabetes (mean 30.7 vs 16.8 years, $p < 0.001$). Patients with type 1 diabetes had the highest HbA_{1c} (mean 9.1% (76 mmol/mol) vs 7.4% (57 mmol/mol) in other types of diabetes, $p < 0.001$), reflecting poorer glycaemic control. Type 2 diabetic patients tended to be more obese compared with the other three subtypes of diabetes (mean BMI [children adjusted to adult values] 33.1 vs 25.1 kg/m², respectively, $p < 0.0001$).

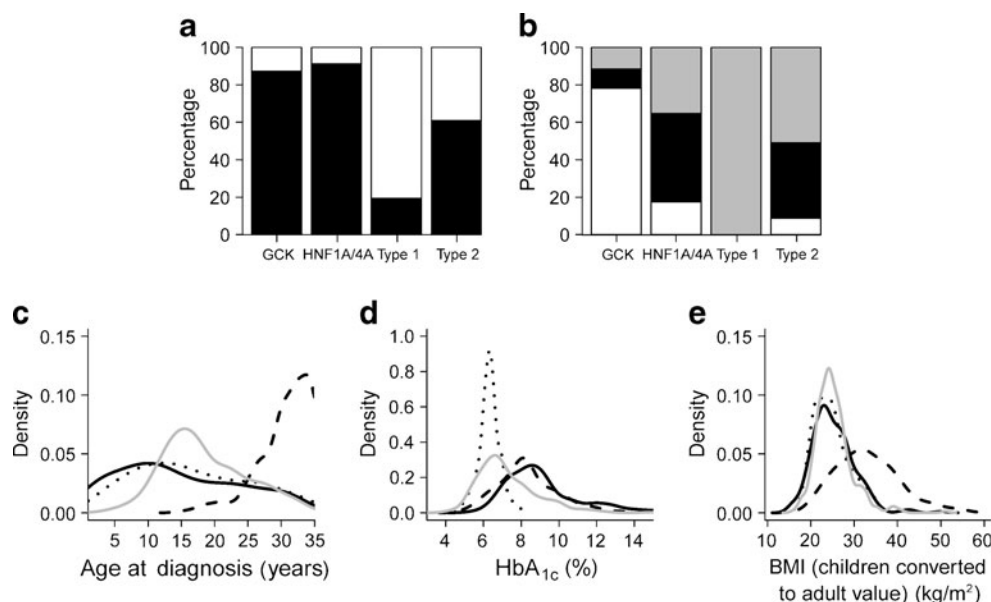
Type 2 diabetes vs MODY models for patients who are not treated with insulin within 6 months of diagnosis

The best model for discriminating type 2 diabetes from MODY was found using logistic regression. Further details of the analysis are provided in ESM Results 1. Table 1 shows the characteristics and associated beta-coefficients and ORs from this model. The strongest predictor of MODY was age at diagnosis, with younger patients being more likely to have MODY (OR for MODY 0.73, [95% CI 0.68, 0.78] for every year increase in age at diagnosis). Other variables associated with higher odds of MODY were being slimmer (as reflected by BMI), having a parent affected with diabetes, not being treated with insulin or OHAs, having better glycaemic control (as reflected by HbA_{1c}), female sex and current age (see Table 1: T2D vs MODY model). It made little difference whether unadjusted BMI, or BMI with child values converted to the adult equivalent, was entered into the model. For ease of use in practice, unadjusted BMI was chosen in the final model.

Figure 2a shows a boxplot of the fitted probabilities for MODY from this logistic regression model for both the MODY and type 2 diabetic patients. The majority of patients with type 2 diabetes have low predicted probabilities for MODY, and the majority of MODY patients have high predicted probabilities. The c-statistic (Fig. 2b) was 0.98, indicating that the model showed excellent discrimination between MODY and type 2 diabetes. Table 2 shows the sensitivities, specificities and corresponding likelihood ratios for MODY using different probability cut-offs in the model. The probability cut-off maximising both sensitivity (92%) and specificity (95%) was found to be 60% (Table 2: T2D vs MODY model).

The model worked well when comparing type 2 diabetes with HNF1A/4A and GCK MODY separately (ROC AUC=

Fig. 1 Patient characteristics. Bar plots showing percentages of (a) parent affected by diabetes (in black) and (b) treatment (diet, white; OHA, black; insulin [\pm OHA], grey). Density plots for (c) age at diagnosis, (d) HbA_{1c} and (e) BMI (with child values converted to adult equivalent). Distributions for the four subtypes of diabetes; type 1, solid black line; type 2, dashed line; GCK MODY, dotted line; HNF1A/4A MODY, solid grey line. To convert values for HbA_{1c} in % into mmol/mol, subtract 2.15 and multiply by 10.929



0.974 and 0.990, respectively). We also examined the subgroup of patients diagnosed between 25 and 35 years with HNF1A/4A MODY, as this is the most difficult group to identify, and found that the model still worked well in separating HNF1A/4A MODY patients from those with type 2 diabetes (ROC AUC=0.915).

Validation The prediction error from jack-knife cross-validation was 5%, suggesting that approximately 95% of cases would be correctly classified given this model. In the external test dataset, the model performed well, with a c-statistic of 0.94. A 60% probability cut-off provided 93% specificity and 87% sensitivity for MODY.

Table 1 Clinical discriminators in a) type 2 diabetes (T2D) vs MODY and b) type 1 diabetes (T1D) vs MODY logistic regression models

Model	Log OR (β)	SE (β)	<i>z</i>	<i>p</i> value	OR (exp[β])	95% CI OR
T2D vs MODY model						
Age at diagnosis (years)	-0.32	0.03	-9.27	<0.0001	0.73	0.68, 0.78
BMI (kg/m ²)	-0.23	0.03	-7.29	<0.0001	0.79	0.74, 0.84
HbA _{1c} (%) ^a	-0.63	0.13	-4.89	<0.0001	0.53	0.41, 0.68
Parent diabetic	1.74	0.42	4.21	<0.0001	5.74	2.61, 13.37
Age (years)	-0.04	0.01	-2.64	0.008	0.97	0.94, 0.99
Insulin or OHA treated	-1.0	0.44	-2.26	0.024	0.37	0.15, 0.87
Sex (male=1, female=2)	0.69	0.34	2.05	0.04	2.00	1.03, 3.93
T1D vs MODY model						
Parent diabetic	3.14	0.34	9.12	<0.0001	23.11	12.10, 46.91
Age (years)	-0.08	0.01	-6.86	<0.0001	0.92	0.90, 0.94
HbA _{1c} (%) ^b	-0.66	0.10	-6.31	<0.0001	0.52	0.42, 0.63
Age at diagnosis (years)	0.10	0.02	4.25	<0.0001	1.11	1.06, 1.16
Sex (male=1, female=2)	1.31	0.35	3.71	0.0002	3.72	1.89, 7.61

Log OR(β), logistic regression β coefficients (log odds ratio for one unit increase in the explanatory variable)

SE(β), standard error for the β coefficient

z, standardised effect size

OR, odds ratio for one unit increase in the explanatory variable (exponential of β)

95% CI OR, 95% confidence interval for the odds ratio

^a Equivalent effect sizes for HbA_{1c} in mmol/mol: β (SE)=-0.06 (0.01); OR (95% CI)=0.94 (0.92, 0.97)

^b Equivalent effect sizes for HbA_{1c} in mmol/mol: β (SE)=-0.06 (0.01); OR (95% CI)=0.94 (0.92, 0.96)

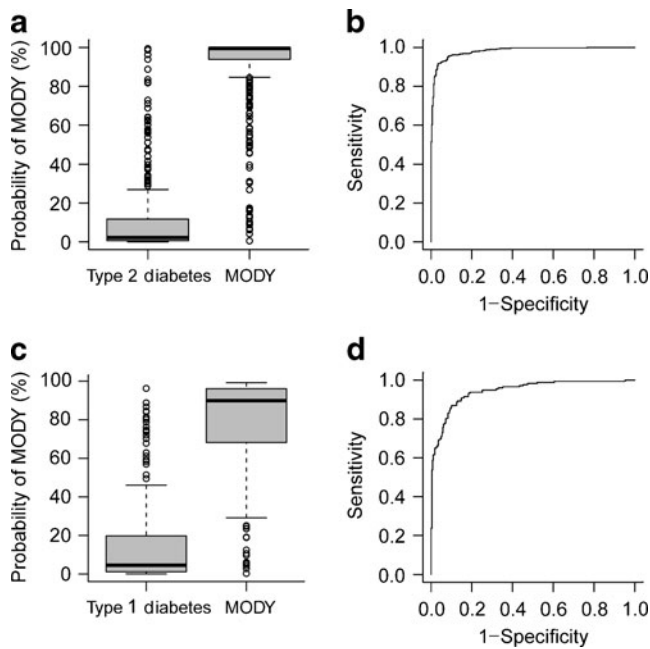


Fig. 2 **a** Boxplot of fitted probabilities for MODY from the type 2 diabetes vs MODY logistic regression model and **(b)** ROC curve showing the discriminative ability of the type 2 diabetes vs MODY logistic regression model; c-statistic=0.98. Similar plots **(c, d)** are shown for the type 1 diabetes vs MODY model; c-statistic=0.95

Type 1 diabetes vs MODY models for patients who are treated with insulin within 6 months of diagnosis

The best model for discriminating type 1 diabetes from MODY was also found using logistic regression. Further details of the analysis are provided in ESM Results 1. The strongest predictor of MODY in this model was having a parent with diabetes, with those with at least one parent affected having odds of MODY approximately 23 times higher (95% CI 12.1, 46.9) than those without a parent affected. Other significant predictors of MODY were lower HbA_{1c}, female sex, older age at diagnosis and current age (see Table 1: T1D vs MODY model).

The model showed good performance with type 1 diabetic patients having low predicted probabilities, and the majority of MODY patients having high predicted probabilities for MODY (Fig. 2c). The c-statistic was 0.95 (Fig. 2d). The sensitivities, specificities and corresponding likelihood ratios for MODY using different probability cut-offs in the model are shown in Table 2. The probability cut-off maximising both sensitivity (87%) and specificity (88%) in this model was found to be 40% (Table 2: T1D vs MODY model).

The model worked well when considering type 1 diabetes with HNF1A/4A and GCK MODY separately (ROC AUC=0.934 and 0.976, respectively), and when comparing HNF1A/4A MODY with type 1 diabetes in only those diagnosed between 25 and 35 years (ROC AUC=0.952).

Validation The prediction error from jack-knife cross-validation was 9.2%, suggesting that approximately 90% of cases would be correctly classified. In the external test dataset, the model performed well, with an ROC AUC of 0.95. A 40% probability cut-off provided 93% specificity and 82% sensitivity for MODY.

Comparison with traditional MODY criteria

Standard criteria for MODY (age at diagnosis <25 years and parent affected with diabetes) were specific, correctly excluding MODY in 543/597 (91%) type 1 diabetes or type 2 diabetes cases, but had lower sensitivity, picking up only 425/594 (72%) MODY cases. Using the optimal cut-offs identified, the prediction models were more sensitive at picking up proven MODY cases (539/594 [91% sensitivity]), with similar specificity (560/597 [94%]), correctly classifying more patients overall (92% vs 81%, $p<0.0001$).

Alternative models

The best linear discriminant and classification tree models are described in the ESM (ESM Results 2, ESM Table 1 and ESM Figure 1).

Positive predictive values (post-test probabilities) for MODY

The positive predictive values and negative predictive values for MODY, adjusting for prior probabilities of 0.7% and 4.6% (for the type 1 diabetes and type 2 diabetes comparisons, respectively) are presented in Table 2.

Compared with the type 2 diabetes vs MODY model the positive predictive values obtained in the type 1 diabetes vs MODY model were more modest, reflecting the lower prior probabilities.

Online clinical prediction calculator

The equations obtained from both logistic regression models were used to produce a web-based version of the models now available at www.diabetesgenes.org. The clinical features of a particular patient diagnosed ≤ 35 years can be entered into this online form. For patients who are treated with insulin within 6 months of diagnosis, the type 1 diabetes vs MODY equation will be applied to their clinical features, otherwise the type 2 diabetes vs MODY equation will be applied, and the positive predictive value for MODY calculated.

Discussion

We have presented models showing excellent discrimination between MODY and both type 1 and type 2 diabetes in

Table 2 Sensitivity, specificity, likelihood ratio and post-test probability for MODY for different cut-offs obtained in a) type 2 diabetes (T2D) vs MODY and b) type 1 diabetes (T1D) vs MODY logistic regression models

Model	Probability cut-off for classifying as MODY from the logistic regression model (%)									
	10	20	30	40	50	60	70	80	90	
T2D vs MODY										
Sensitivity/ specificity (%)	99/73	97/82	97/86	96/91	94/92	92/95	90/97	86/98	80/99	
LR+ for MODY ^a (95% CI)	3.7 (3.1, 4.4)	5.5 (4.4, 7.0)	6.7 (5.1, 8.8)	10.2 (7.2, 14.3)	11.6 (8.0, 16.7)	17.3 (10.9, 27.5)	28.6 (15.5, 52.7)	34.4 (17.4, 68.3)	63.9 (24.1, 169.4)	
PPV for MODY ^b (%)	15.1	21.0	24.4	32.9	35.8	45.5	58.0	62.4	75.5	
LR- for MODY ^c (95% CI)	0.02 (0.01, 0.04)	0.04 (0.02, 0.06)	0.04 (0.02, 0.07)	0.05 (0.03, 0.07)	0.06 (0.04, 0.09)	0.08 (0.06, 0.11)	0.11 (0.08, 0.13)	0.14 (0.11, 0.18)	0.20 (0.17, 0.24)	
NPV for MODY ^d (%)	99.9	99.8	99.8	99.8	99.7	99.6	99.5	99.3	99.0	
T1D vs MODY										
Sensitivity/ specificity (%)	96/65	94/76	91/85	87/88	83/91	80/93	73/94	66/97	50/99.6	
LR+ for MODY ^a (95% CI)	2.8 (2.4, 3.3)	3.8 (3.1, 4.7)	5.9 (4.5, 7.8)	7.3 (5.3, 10.2)	9.6 (6.5, 14.2)	11.1 (7.2, 17.0)	12.7 (7.8, 20.6)	20.4 (10.6, 39.2)	138.2 (19.4, 983.3)	
PPV for MODY ^b (%)	1.9	2.6	4.0	4.9	6.4	7.2	8.2	12.6	49.4	
LR- for MODY ^c (95% CI)	0.06 (0.03, 0.13)	0.08 (0.05, 0.15)	0.11 (0.07, 0.17)	0.15 (0.10, 0.22)	0.19 (0.13, 0.26)	0.22 (0.16, 0.30)	0.29 (0.23, 0.37)	0.35 (0.28, 0.43)	0.50 (0.44, 0.58)	
NPV for MODY ^d (%)	99.96	99.94	99.92	99.90	99.87	99.85	99.80	99.75	99.65	

^a Positive likelihood ratio (LR) for MODY = sensitivity/(1 - specificity)^b Positive predictive value (PPV; post-test probability) for MODY = (sensitivity × prevalence) / [(sensitivity × prevalence) + ((1 - specificity) × [1 - prevalence])], assuming prevalences for MODY of 4.6% for the T2D vs MODY and 0.7% for T1D vs MODY models, respectively^c Negative likelihood ratio for MODY = (1 - sensitivity) / specificity^d Negative predictive value (NPV) for MODY = [specificity × (1 - prevalence)] / [(specificity × [1 - prevalence]) + ((1 - sensitivity) × prevalence)]

patients diagnosed ≤ 35 years, with c-statistics > 0.94 in both training and test datasets, and cross-validated prediction errors of $< 10\%$. The models show considerable improvement in prediction of MODY compared with traditional clinical criteria. The models are validated, and provide post-test probabilities for MODY that would be helpful in clinical practice, particularly as supported by an online calculator. Probabilities are produced, rather than classifications, allowing decisions regarding testing to vary in light of other information.

Genetic testing is highly specific and sensitive and represents the gold standard for diagnosing MODY. However, the cost limits its use in all individuals. The decision on whether to test is a clinical judgement that depends on the likelihood and impact of a positive result, weighted against the expense of the test. The current pick-up rate based on UK referrals in the Exeter genetics diagnostic laboratory is 27% [3], which is based on pre-selection of patients by healthcare professionals in line with best-practice guidelines [14]. We would advise that this model should be used in all patients diagnosed with diabetes under the age of 35. In patients who are not treated with insulin within 6 months of diagnosis, post-test probabilities of $> 25\%$ would be appropriate to trigger genetic testing. In patients who are treated with insulin within 6 months of diagnosis, a lower post-test probability of $> 10\%$ could be considered appropriate as the impact of finding a mutation is greater than for a non-insulin-treated patient, both financially, through saved treatment costs, and in terms of improved quality of life. In these cases, further testing of C-peptide and pancreatic autoantibodies could be carried out prior to genetic testing, with a positive C-peptide and negative autoantibody result being strongly suggestive of MODY compared with type 1 diabetes [19, 20]. As the cost of genetic testing decreases, the threshold at which to refer patients will lower. A full health-economic model would be required to formally explore the relative trade-offs.

The pre-test probability of MODY we used is likely to be an underestimate, particularly in the paediatric age group. In a study of white non-insulin-treated patients under 17, it was proposed that the probability of MODY was similar to that of type 2 diabetes [21], much higher than 4.6% as used in our study. In all studies to date, prevalence estimates have been restricted by initial selection on clinical criteria before genetic testing, and the proportion of patients missed has not been quantified. Small increases in pre-test probabilities could make a considerable difference to the post-test probabilities.

We have provided internal and external validation of the model, but further testing is needed. The model should be validated in different settings and populations, particularly as our dataset comprised solely white Europeans. The performance of the model in other populations could vary considerably as the incidence of type 2 diabetes in adolescents and young adults is more common in high-risk ethnic groups [22]. Further validation on a community-based

population cohort would remove the biases associated with being referred to the diagnostic service, and would elucidate whether the female predominance is genuine.

We cannot exclude the possibility that patients in the type 1 diabetes or type 2 diabetes groups did not have MODY, as genetic testing was not carried out on them all. It is also possible that a few of the confirmed MODY patients may also have coincidental type 1 or type 2 diabetes. These misclassified patients will be a tiny minority of cases and would have little impact on the final model.

As genetic testing becomes cheaper, its use in diagnosis will be more commonplace. One consequence of this will be that more variants will be identified with greater uncertainty as to whether they are pathogenic mutations or rare polymorphisms of no clinical significance. Probability models based on clinical criteria, such as those presented here, may then have a different but equally important role in producing prior probabilities for MODY to inform the likelihood of pathogenicity of a novel mutation.

Incorporating other characteristics into the model could improve its diagnostic ability. Only basic clinical characteristics were used in these models and data were limited to information available for all participants. Some important clinical features known to be indicative of MODY, type 1 diabetes and type 2 diabetes were not included. Pancreatic autoantibodies [23] and measures of endogenous insulin secretion, such as persistent C-peptide [24], are highly sensitive and specific biomarkers of type 1 diabetes, which have been shown to discriminate well between MODY and type 1 diabetes [19, 20]. Young-onset type 2 diabetes is more common in non-white populations and is often characterised by signs of insulin resistance, such as acanthosis nigricans [22], which is rare in MODY. HNF1A MODY patients show sensitivity to sulfonylureas [4, 5], have a low renal threshold for glucose [25], and lower high-sensitivity C-reactive protein (hsCRP) levels [13]. A small glucose increment in an OGTT is predictive of a diagnosis of GCK MODY [26]. Information on these additional factors is not always collected. The advantage of the clinical criteria we have used is that it should be routinely available for all patients. We have already shown excellent discrimination (ROC AUC > 0.94) using clinical characteristics alone. The next important step will be to assess the additional value of other biomarkers, which would be most useful in aiding decisions in insulin-treated patients. Provided that they are independent, it is possible to account for other information by combining likelihood ratios for any additional test with that produced by the model.

In conclusion, we have produced a clinical prediction model that shows good discrimination between MODY and the more common type 1 diabetes and type 2 diabetes in patients diagnosed under the age of 35. This should provide a useful aid for selecting patients for diagnostic genetic testing for MODY who may benefit from improved treatment and management if a mutation is identified.

Acknowledgements We gratefully acknowledge the Warren 2 consortium, K. Colclough of the Royal Devon and Exeter NHS Foundation Trust, and O. Rubio-Cabezas, A. Steele, A. Hill, M. Hudson and other members of the Exeter Clinical Research team at the Peninsula Medical School, who have collected samples and provided data. The authors thank E. Pearson and R. McAlpine of Tayside NHS Trust for providing routine prevalence data.

Funding This study was funded by the European Community FP7 programme CEED3 (HEALTH-F2-2008-223211) and the Health Innovation Challenge Fund (HICF-1009-0041). BS, SE, and ATH are supported by the National Institute of Health Research (NIHR) Peninsula Clinical Research Facility, University of Exeter. CH is supported in this work by the PenCLAHRC programme for the NIHR.

Duality of interest The authors declare that there is no duality of interest associated with this manuscript. The article presents independent research commissioned by the NIHR and the views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Contribution statement BS designed the study, collated all data, carried out all analysis and drafted the manuscript. AH conceived of the idea and provided supervision and critical revision of the manuscript. TM assisted with study design and interpretation of analysis and provided critical revision of the manuscript. SE assisted with acquisition of data, interpretation of data, and provided critical revision of the manuscript. MC provided supervision on design of the regression models, advised on statistical analysis and provided critical revision of the manuscript. CH assisted with interpretation of analysis and critical revision of the manuscript. All authors approved the final version of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Tattersall RB (1974) Mild familial diabetes with dominant inheritance. *Q J Med* 43:339–357
2. Tattersall RB, Fajans SS (1975) A difference between the inheritance of classical juvenile-onset and maturity-onset type diabetes of young people. *Diabetes* 24:44–53
3. Shields BM, Hicks S, Shepherd MH, Colclough K, Hattersley AT, Ellard S (2010) Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia* 53:2504–2508
4. Pearson ER, Starkey BJ, Powell RJ, Gribble FM, Clark PM, Hattersley AT (2003) Genetic cause of hyperglycaemia and response to treatment in diabetes. *Lancet* 362:1275–1281
5. Shepherd M, Shields B, Ellard S, Rubio-Cabezas O, Hattersley AT (2009) A genetic diagnosis of HNF1A diabetes alters treatment and improves glycaemic control in the majority of insulin-treated patients. *Diabet Med* 26:437–441
6. Pearson ER, Velho G, Clark P et al (2001) Beta-cell genes and diabetes: quantitative and qualitative differences in the pathophysiology of hepatic nuclear factor-1alpha and glucokinase mutations. *Diabetes* 50(Suppl 1):S101–S107
7. Froguel P, Zouali H, Vionnet N et al (1993) Familial hyperglycaemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus. *N Engl J Med* 328:697–702
8. Schober E, Rami B, Grabert M et al (2009) Phenotypical aspects of maturity-onset diabetes of the young (MODY diabetes) in comparison with Type 2 diabetes mellitus (T2DM) in children and adolescents: experience from a large multicentre database. *Diabet Med* 26:466–473
9. Galler A, Stange T, Muller G et al (2010) Incidence of childhood diabetes in children aged less than 15 years and its clinical and metabolic characteristics at the time of diagnosis: data from the childhood diabetes registry of Saxony, Germany. *Horm Res Paediatr* 74:285–291
10. Eide SA, Raeder H, Johansson S et al (2008) Prevalence of HNF1A (MODY3) mutations in a Norwegian population (the HUNT2 Study). *Diabet Med* 25:775–781
11. Ledermann HM (1995) Maturity-onset diabetes of the young (MODY) at least ten times more common in Europe than previously assumed? *Diabetologia* 38:1482
12. Fajans SS, Bell GI, Bowden DW, Halter JB, Polonsky KS (1994) Maturity-onset diabetes of the young. *Life Sci* 55:413–422
13. Owen KR, Thanabalasingham G, James TJ et al (2010) Assessment of high-sensitivity C-reactive protein levels as diagnostic discriminator of maturity-onset diabetes of the young due to HNF1A mutations. *Diabetes Care* 33:1919–1924
14. Ellard S, Bellanne-Chantelot C, Hattersley AT (2008) Best practice guidelines for the molecular genetic diagnosis of maturity-onset diabetes of the young. *Diabetologia* 51:546–553
15. Rodriguez-Bigas MA, Boland CR, Hamilton SR et al (1997) A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *J Natl Cancer Inst* 89:1758–1762
16. Wells PS, Anderson DR, Rodger M et al (2000) Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpliRED D-dimer. *Thromb Haemost* 83:416–420
17. Royal College of General Practitioners and NHS Diabetes (2011) Coding, Classification and Diagnosis of Diabetes. Available from www.diabetes.nhs.uk/our_work_areas/classification_of_diabetes/2011, accessed 16 Aug 2011
18. Freeman JV, Cole TJ, Chinn S, Jones PR, White EM, Preece MA (1995) Cross sectional stature and weight reference curves for the UK, 1990. *Arch Dis Child* 73:17–24
19. Besser RE, Shepherd MH, McDonald TJ et al (2011) Urinary C-peptide creatinine ratio is a practical outpatient tool for identifying hepatocyte nuclear factor 1 α /hepatocyte nuclear factor 4- α maturity-onset diabetes of the young from long-duration type 1 diabetes. *Diabetes Care* 34:286–291
20. McDonald TJ, Colclough K, Brown R et al (2011) Islet autoantibodies can discriminate maturity-onset diabetes of the young (MODY) from type 1 diabetes. *Diabet Med* 28(9):1028–1033
21. Ehtisham S, Hattersley AT, Dunger DB, Barrett TG (2004) First UK survey of paediatric type 2 diabetes and MODY. *Arch Dis Child* 89:526–529
22. Rosenbloom AL, Joe JR, Young RS, Winter WE (1999) Emerging epidemic of type 2 diabetes in youth. *Diabetes Care* 22:345–354
23. Barker JM (2006) Clinical review: type 1 diabetes-associated autoimmunity: natural history, genetic associations, and screening. *J Clin Endocrinol Metab* 91:1210–1217
24. Koskinen P, Viikari J, Irjala K, Kaihola HL, Seppala P (1986) Plasma and urinary C-peptide in the classification of adult diabetics. *Scand J Clin Lab Invest* 46:655–663
25. Menzel R, Kaisaki PJ, Rjasanowski I, Heinke P, Kerner W, Menzel S (1998) A low renal threshold for glucose in diabetic patients with a mutation in the hepatocyte nuclear factor-1alpha (HNF-1alpha) gene. *Diabet Med* 15:816–820
26. Stride A, Vaxillaire M, Tuomi T et al (2002) The genetic abnormality in the beta cell determines the response to an oral glucose load. *Diabetologia* 45:427–435