File S1: Populated TRIPOD+AI checklist

Who should have an HbA1c? Development and external validation of a dysglycaemia prediction model using only routinely available data

This study both develops and evaluates a prediction model, so every item of the TRIPOD+AI checklist applies. The topic column is an abbreviated label; the full wording of each item is in Table 2 of the TRIPOD+AI statement (Collins GS, Moons KGM, Dhiman P, et al. BMJ 2024;385:e078378) and the official fillable template is available at www.tripod-statement.org. Items marked not applicable are explained in the same row, as the statement requires.

Item	Topic	Where reported in this manuscript	Status
1	Title	Title	Reported
2	Abstract	Abstract (Background / Methods / Results / Conclusions)	Reported
3a	Background: context, rationale, existing models	Background, paragraphs 1-6	Reported
3b	Target population, intended purpose and users	Background, paragraphs 2-3; Discussion, implementation paragraph	Reported
3c	Known health inequalities between sociodemographic groups	Background, paragraph 1	Reported
6a	Study setting, number and location of centres	Methods, Data source and participants (national multistage probability survey with mobile examination centres; individual centre locations are not released by NHANES)	Reported
5a	Sources of data, rationale, representativeness	Methods, Data source and participants	Reported
5b	Dates of participant data	Methods, Data source and participants; Abstract	Reported
6a	Study setting, number and location of centres	Methods, Data source and participants	Reported
6b	Eligibility criteria	Methods, Data source and participants; Fig. 1	Reported
6c	Treatments received and how handled	Methods, Predictors (antihypertensive and lipid-lowering medication are predictors)	Reported
7	Data pre-processing and quality checking	Methods, Predictors and Statistical analysis; Additional file 1, Table S2	Reported
8a	Outcome definition, how and when assessed, rationale	Methods, Outcome	Reported
8b	Qualifications of outcome	Not applicable: the outcome	Not applicable

	assessors if subjective	is a laboratory HbA1c measured by a standardised NHANES protocol	
8c	Blinding of outcome assessment	Not applicable: laboratory measurement, no assessor judgement	Not applicable
9a	Choice of initial predictors and any pre-selection	Methods, Predictors	Reported
9b	Definition of all predictors, how and when measured	Methods, Predictors; Additional file 1, Tables S1 and S2	Reported
9c	Qualifications of predictor assessors if subjective	Not applicable: measured or recorded by standardised NHANES protocol	Not applicable
10	How study size was arrived at	Methods, Outcome (events per candidate predictor, with reference)	Reported
11	Handling of missing data	Methods, Statistical analysis; Additional file 1, Table S2	Reported
12a	How data were used, partitioning	Methods, Data source and participants and Statistical analysis	Reported
12b	Handling of predictors in the analysis	Methods, Statistical analysis (restricted cubic splines, standardisation)	Reported
12c	Model type, rationale, building steps, tuning, internal validation	Methods, Statistical analysis	Reported
12d	Heterogeneity across clusters	Not applicable: a single national survey, no centre-level clustering modelled beyond the survey design	Not applicable
12e	Measures and plots used to evaluate performance	Methods, Statistical analysis	Reported
12f	Model updating arising from evaluation	Not applicable: no updating or recalibration was performed	Not applicable
12g	How model predictions were calculated for evaluation	Methods, Statistical analysis; Additional file 1, Table S1 and reference implementation	Reported
13	Class imbalance methods	Methods, Statistical analysis: none used, with rationale	Reported
14	Approaches to model fairness	Methods, Statistical analysis; Results, Subgroup performance; Discussion	Reported
15	Model output and any classification thresholds	Methods, Statistical analysis; Results, Clinical utility; Table 4	Reported
16	Differences between development and evaluation data	Methods, Predictors and Comparator; Table 1	Reported
17	Ethics approval and consent	Methods, Ethics; Declarations	Reported
18a	Funding and role of funders	Declarations, Funding	Reported
18b	Conflicts of interest	Declarations, Competing interests	Reported
18c	Protocol availability	Declarations, Protocol and registration	Reported
18d	Registration	Declarations, Protocol and registration: not registered	Reported

18e	Data availability	Declarations, Availability of data and materials	Reported
18f	Code availability	Declarations, Availability of data and materials	Reported
19	Patient and public involvement	Declarations, Patient and public involvement: none	Reported
20a	Flow of participants, with and without the outcome	Results, Participants; Fig. 1	Reported
20b	Characteristics, key dates, predictors, events, missing data; differences across demographic groups	Table 1; Additional file 1, Table S2; Results, Subgroup performance	Reported
20c	Comparison of predictor distributions, development versus evaluation	Table 1 (both cycles side by side)	Reported
21	Participants and events in each analysis	Results, Participants; Tables 2 and 3; Additional file 1, Table S3	Reported
22	Full model details enabling prediction in new individuals	Additional file 1, Table S1; deposited reference implementation	Reported
23a	Performance with confidence intervals, including key subgroups	Table 2; Results, Subgroup performance; Additional file 1, Table S5; Figs. 2 and 3	Reported
23b	Heterogeneity in performance across clusters	Not applicable: not examined; see item 12d	Not applicable
24	Results of any model updating	Not applicable: no updating performed	Not applicable
25	Overall interpretation, including fairness	Discussion	Reported
26	Limitations and their effects	Discussion, limitations paragraphs	Reported
27a	Handling of poor quality or unavailable input data at implementation	Discussion, implementation paragraph	Reported
27b	User interaction and expertise required	Discussion, implementation paragraph	Reported
27c	Next steps for future research, applicability and generalisability	Discussion, final paragraph	Reported

Of the 52 subitems, 45 are reported and 7 are not applicable to this study. No subitem is unreported.