TRIPOD+AI checklist for the reporting of prediction model studies

Section/ Topic	Item No	Development/ Evaluation*	Checklist Item	Reported on Page Number/Line Number	Reported on Section/Paragraph
Title					
Title	1	D;E	Identify the study as developing or evaluating the performance of a multivariable prediction model, the target population, and the outcome to be predicted	1/1-3	Title
Abstract					
Abstract	2	D;E	See TRIPOD+AI for Abstracts checklist	2-3/21-56	Abstract/1-4
Introduction					
Background	3a	D;E	Explain the healthcare context (including whether diagnostic or prognostic) and rationale for developing or evaluating the prediction model, including references to existing models	4-6/64-121	Introduction/1-4
	3b	D;E	Describe the target population and the intended purpose of the prediction model in the context of the care pathway, including its intended users (eg, healthcare professionals, patients, public)	6-7/122-140	Introduction/5
	3c	D;E	Describe any known health inequalities between sociodemographic groups	4/64-74	Introduction/1
Objectives	4	D;E	Specify the study objectives, including whether the study describes the development or validation of a prediction model (or both)	6-7/122-140	Introduction/5
Methods					
Data	5a	D;E	Describe the sources of data separately for the development and evaluation datasets (eg, randomised trial, cohort, routine care or registry data), the rationale for using these data, and representativeness of the data	7/145-149	Methods/1
	5b	D;E	Specify the dates of the collected participant data, including start and end of participant accrual; and, if applicable, end of follow-up	7/145-149	Methods/1
Participants	6a	D;E	Specify key elements of the study setting (eg, primary care, secondary care, general population) including the number and location of centres	7-8/150-163	Methods/2
	6b	D;E	Describe the eligibility criteria for study participants	7-8/150-163	Methods/2
	6c	D;E	Give details of any treatments received, and how they were handled during model development or evaluation, if relevant	8-9/164-174	Methods/3

Data preparation	7	D;E	Describe any data pre-processing and quality checking, including whether this was similar across relevant sociodemographic groups	7-9/150-174	Methods/2-3
Outcome	8a	D;E	Clearly define the outcome that is being predicted and the time horizon, including how and when assessed, the rationale for choosing this outcome, and whether the method of outcome assessment is consistent across sociodemographic groups	8-9/164-174	Methods/3
	8b	D;E	If outcome assessment requires subjective interpretation, describe the qualifications and demographic characteristics of the outcome assessors	8-9/164-174	Methods/3
	8c	D;E	Report any actions to blind assessment of the outcome to be predicted	8-9/164-174	Methods/3
Predictors	9a	D	Describe the choice of initial predictors (eg, literature, previous models, all available predictors) and any pre-selection of predictors before model building	9-10/181-193	Methods/5
	9b	D;E	Clearly define all predictors, including how and when they were measured (and any actions to blind assessment of predictors for the outcome and other predictors)	10/194-201	Methods/6
	9c	D;E	If predictor measurement requires subjective interpretation, describe the qualifications and demographic characteristics of the predictor assessors	7-9/150-174	Methods/2-3
Sample size	10	D;E	Explain how the study size was arrived at (separately for development and evaluation), and justify that the study size was sufficient to answer the research question. Include details of any sample size calculation	7-9/150-174	Methods/2-3
Missing data	11	D;E	Describe how missing data were handled. Provide reasons for omitting any data	N/A. No missing data.	N/A. No missing data.
Analytical methods	12a	D	Describe how the data were used (eg, for development and evaluation of model performance) in the analysis, including whether the data were partitioned, considering any sample size requirements	10-12/202-239	Methods/7-10
	12b	D	Depending on the type of model, describe how predictors were handled in the analyses (functional form, rescaling, transformation, or any standardisation)	11/210-222	Methods/8-9
	12c	D	Specify the type of model, rationale[†], all model building steps, including any hyperparameter tuning, and method for internal validation	13-16/240-319	Methods/11-16
	12d	D;E	Describe if and how any heterogeneity in estimates of model parameter values and model performance was handled and quantified across clusters (eg, hospitals, countries). See TRIPOD-Cluster for additional considerations[‡]	N/A. No acrossing clusters.	N/A. No acrossing clusters.
	12e	D;E	Specify all measures and plots used (and their rationale) to evaluate model performance (eg, discrimination, calibration, clinical utility) and, if relevant, to compare multiple models	16-19/320-369	Methods/17-20

	12f	E	Describe any model updating (eg, recalibration) arising from the model evaluation, either overall or for particular sociodemographic groups or settings	14-16/262-319	Methods/12-16
	12g	E	For model evaluation, describe how the model predictions were calculated (eg, formula, code, object, application programming interface)	17-19/333-369	Methods/18-20
Class imbalance	13	D;E	If class imbalance methods were used, state why and how this was done, and any subsequent methods to recalibrate the model or the model predictions	N/A. No class imbalance.	N/A. No class imbalance.
Fairness	14	D;E	Describe any approaches that were used to address model fairness and their rationale	14-15/274-300	Methods/13-14
Model output	15	D	Specify the output of the prediction model (eg, probabilities, classification). Provide details and rationale for any classification and how the thresholds were identified	19-20/377-387	Methods/22-23
Training versus evaluation	16	D;E	Identify any differences between the development and evaluation data in healthcare setting, eligibility criteria, outcome, and predictors	N/A. No differences.	N/A. No differences.
Ethical approval	17	D;E	Name the institutional research board or ethics committee that approved the study and describe the participant informed consent or the ethics committee waiver of informed consent	7/145-149	Methods/1
Open science					
Funding	18a	D;E	Give the source of funding and the role of the funders for the present study	35/658-663	Acknowledgments
Conflicts of interest	18b	D;E	Declare any conflicts of interest and financial disclosures for all authors	N/A. No conflicts.	N/A. No conflicts.
Protocol	18c	D;E	Indicate where the study protocol can be accessed or state that a protocol was not prepared	36/671-672	Footnote/3
Registration	18d	D;E	Provide registration information for the study, including register name and registration number, or state that the study was not registered	36/668-670	Footnote/1
Data sharing	18e	D;E	Provide details of the availability of the study data	36/671-672	Footnote/3
Code sharing	18f	D;E	Provide details of the availability of the analytical code[§]	36/671-672	Footnote/3
Patient and public involvement					
Patient and public involvement	19	D;E	Provide details of any patient and public involvement during the design, conduct, reporting, interpretation, or dissemination of the study or state no involvement	7/145-149	Methods/1

Result					
Participants	20a	D;E	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful	Figure 1	Methods/1-2
	20b	D;E	Report the characteristics overall and, where applicable, for each data source or setting, including the key dates, key predictors (including demographics), treatments received, sample size, number of outcome events, follow-up time, and amount of missing data. A table may be helpful. Report any differences across key demographic groups	Table 2-3	Table 2-3
	20c	E	For model evaluation, show a comparison with the development data of the distribution of important predictors (demographics, predictors, and outcome)	Table 3	Table 3
Model development	21	D;E	Specify the number of participants and outcome events in each analysis (eg, for model development, hyperparameter tuning, model evaluation)	Table 4-7	Table 4-7
Model specification	22	D	Provide details of the full prediction model (eg, formula, code, object, application programming interface) to allow predictions in new individuals and to enable third party evaluation and implementation, including any restrictions to access or reuse (eg, freely available, proprietary)[¶]	Supplementary	Supplementary
Model performance	23a	D;E	Report model performance estimates with confidence intervals, including for any key subgroups (eg, sociodemographic). Consider plots to aid presentation	Table 3,5-7	Table 3,5-7
	23b	D;E	If examined, report results of any heterogeneity in model performance across clusters. See TRIPOD-Cluster for additional details[‡]	N/A. No acrossing clusters.	N/A. No acrossing clusters.
Model updating	24	E	Report the results from any model updating, including the updated model and subsequent performance	Table 6-7	Table 6-7
Discussion					
Interpretation	25	D;E	Give an overall interpretation of the main results, including issues of fairness in the context of the objectives and previous studies	29-30/525-548	Discussion/1
Limitations	26	D;E	Discuss any limitations of the study (such as a non-representative sample, sample size, overfitting, missing data) and their effects on any biases, statistical uncertainty, and generalisability	34/623-637	Discussion/6
Usability of the model in the context of current care	27a	D	Describe how poor quality or unavailable input data (eg, predictor values) should be assessed and handled when implementing the prediction model	31-32/579-602	Discussion/3
	27b	D	Specify whether users will be required to interact in the handling of the input data or use of the model, and what level of expertise is required of users	30-33/549-622	Discussion/2-5
	27c	D;E	Discuss any next steps for future research, with a specific view to applicability and generalisability of the model	34/623-637	Discussion/6

TRIPOD=Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; AI=artificial intelligence.

* D=items relevant only to the development of a prediction model; E=items relating solely to the evaluation of a prediction model; D;E=items applicable to both the development and evaluation of a prediction model.

† Separately for all model building approaches.

‡ TRIPOD-Cluster is a checklist of reporting recommendations for studies developing or validating models that explicitly account for clustering or explore heterogeneity in model performance (eg, at different hospitals or centres).[1-2]

§ Relates to the analysis code, for example, any data cleaning, feature engineering, model building, and evaluation.

¶ Relates to the code to implement the model to get estimates of risk for a new individual.

1. Debray TPA, Collins GS, Riley RD, et al. Transparent reporting of multivariable prediction models developed or validated using clustered data: TRIPOD-Cluster checklist. BMJ 2023;380:e071018. doi:10.1136/bmj-2022-071018

2. Debray TPA, Collins GS, Riley RD, et al. Transparent reporting of multivariable prediction models developed or validated using clustered data (TRIPOD-Cluster): explanation and elaboration. BMJ 2023;380:e071058. doi:10.1136/bmj-2022-071058

Essential items to include for the reporting of prediction model studies in a journal or conference abstract (TRIPOD+AI for Abstracts*)

Section and item	Item No	Checklist Item	Reported on Page Number/Line Number	Reported on Section/Paragraph
Title	1	Identify the study as developing or evaluating the performance of a multivariable prediction model, the target population, and the outcome to be predicted	1/1-3	Title
Background	2	Provide a brief explanation of the healthcare context and rationale for developing or evaluating the performance of all models	2/21-25	Abstract/1
Objectives	3	Specify the study objectives, including whether the study describes model development, evaluation, or both	2/26-30	Abstract/1
Methods	4	Describe the sources of data	2/31-33	Abstract/2
	5	Describe the eligibility criteria and setting where the data were collected	2/33-36	Abstract/2
	6	Specify the outcome to be predicted by the model, including time horizon of predictions in case of prognostic models	2-3/36-42	Abstract/2
	7	Specify the type of model, a summary of the model-building steps, and the method for internal validation†	2-3/36-43	Abstract/2
	8	Specify the measures used to assess model performance (eg, discrimination, calibration, clinical utility)	3/42-43	Abstract/2

Results	9	Report the number of participants and outcome events	3/44-48	Abstract/3
	10	Summarise the predictors in the final model[†]	3/48-50	Abstract/3
	11	Report model performance estimates (with confidence intervals)	3/50-53	Abstract/3
Discussion	12	Give an overall interpretation of the main results	3/54-56	Abstract/4
Registration	13	Give the registration number and name of the registry or repository	36/668-670	Footnote/1

TRIPOD=Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; AI=artificial intelligence.

* This checklist is based on the TRIPOD for Abstracts statement published in 2020,[1] but has been revised and updated for consistency with the TRIPOD+AI statement.

[†] Relevant only to studies describing the development of a prediction model.

1. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1-73. doi:10.7326/M14-0698

Article information: https://dx.doi.org/10.21037/qims-24-1050

*As the checklist was provided upon initial submission, the page number/line number reported may be changed due to copyediting and may not be referable in the published version. In this case, the section/paragraph may be used as an alternative reference.