TRIPOD Checklist: Prediction Model Development and Validation

Section	Item		Checklist description	Reported on Page Number/Line Number	Reported on Section/Paragraph
Title and abstract					
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.		
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.		
Introduction					
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.		
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.		
Methods					
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, ifapplicable.		
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.		
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.		
	5b	D;V	Describe eligibility criteria for participants.		
	5c	D;V	Give details of treatments received, if relevant.		
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.		
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.		
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.		
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.		
Sample size	8	D;V	Explain how the study size was arrived at.		

Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.		
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.		
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.		
	10c	V	For validation, describe how the predictions were calculated.		
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.		
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.		
Risk groups	11	D;V	Provide details on how risk groups were created, if done.		
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.		
Results					
Participants	13a	D;V	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.		
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.		
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).		
Model development	14a	D	Specify the number of participants and outcome events in each analysis.		
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.		
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).		
	15b	D	Explain how to the use the prediction model.		
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.		
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).		
Discussion					
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).		

Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.		
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence.		
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.		
Other information					
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.		
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.		

* Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

Materials Design Analysis Reporting (MDAR)
Checklist for Authors

The MDAR framework establishes a minimum set of requirements in transparent reporting applicable to studies in the life sciences (see Statement of Task: doi:10.31222/osf.io/9sm4x.). The MDAR checklist is a tool for authors, editors and others seeking to adopt the MDAR framework for transparent reporting in manuscripts and other outputs. Please refer to the MDAR Elaboration Document for additional context for the MDAR framework.

Materials

Antibodies	Yes (indicate where provided:	n/a
For commercial reagents, provide supplier name, catalogue number and RRID, if available.		n/a No antibodies were used.

Cell materials	Yes (indicate where provided:	n/a
Cell lines: Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID	Methods/ 8 Paragraph Page 7-8/line 230-243	
Primary cultures: Provide species, strain, sex of origin, genetic modification status.		n/a We used cancer cell lines.

Experimental animals	Yes (indicate where provided:	n/a
Laboratory animals: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog		n/a No animals were used.
Animal observed in or captured from the field: Provide species, sex and age where possible		n/a No animals were used.
Model organisms: Provide Accession number in repository (where relevant) **OR** RRID		n/a No animals were used.

Plants and microbes	Yes (indicate where provided:	n/a
Plants: provide species and strain, unique accession number if available, and source (including location for collected wild specimens)		n/a No plants were used.
Microbes: provide species and strain, unique accession number if available, and source		n/a No microbes were used.

Human research participants	Yes (indicate where provided:	n/a
Identify authority granting ethics approval (IRB or equivalent committee(s), provide reference number for approval.		n/a No human research participants were used.
Provide statement confirming informed consent obtained from study participants.		n/a No human research participants were used.
Report on age and sex for all study participants.		n/a No human research participants were used.

Design

Study protocol	Yes (indicate where provided:	n/a
For clinical trials, provide the trial registration number **OR** cite DOI in manuscript.		n/a No human research participants were used.

Laboratory protocol	Yes (indicate where provided:	n/a
Provide DOI or other citation details if detailed step-by-step protocols are available.	Methods/ 8-12 Paragraph Page 7-9/line 230-280	

Experimental study design (statistics details)	Yes (indicate where provided:	n/a
State whether and how the following have been done**, or** if they were not carried out.		No human research participants were used.
Sample size determination		n/a
Randomisation		n/a
Blinding		n/a
Inclusion/exclusion criteria		n/a

Sample definition and in-laboratory replication	Yes (indicate where provided:	n/a
State number of times the experiment was replicated in laboratory	Methods/12 Paragraph Page 9/line 277-280	
Define whether data describe technical or biological replicates	Methods/12 Paragraph Page 9/line 277-280	

Ethics	Yes (indicate where provided:	n/a
Studies involving human participants: State details of authority granting ethics approval (IRB or equivalent committee(s), provide reference number for approval		n/a No human research participants were used.
Studies involving experimental animals: State details of authority granting ethics approval (IRB or equivalent committee(s), provide reference number		n/a No animals were used.
Studies involving specimen and field samples: State if relevant permits obtained, provide details of authority approving study; if none were required, explain why.		n/a No specimen and field samples were used.

Dual Use Research of Concern (DURC)	Yes (indicate where provided:	n/a
If study is subject to dual use research of concern, state the authority granting approval and reference number for the regulatory approval		n/a The study wasn't subject to dual use research of concern

Analysis

Attrition	Yes (indicate where provided:	n/a
State if sample or data point from the analysis is excluded, and whether the criteria for exclusion were determined and specified in advance.	Methods/1 Paragraph Page 5-6/line 163-174	

Statistics	Yes (indicate where provided:	n/a
Describe statistical tests used and justify choice of tests.	Methods/12 Paragraph Page 9/line 277-280	

Data Availability	Yes (indicate where provided:	n/a
State whether newly created datasets are available, including protocols for access or restriction on access.		n/a All data analyzed in the manuscript were from the open database TCGA, and thus there weren't new data to be shared.
If data are publicly available, provide accession number in repository or DOI or URL.		n/a All data analyzed in the manuscript were from the open database TCGA, and thus there weren't new data to be shared.
If publicly available data are reused, provide accession number in repository or DOI or URL, where possible.		n/a All data analyzed in the manuscript were from the open database TCGA, and thus there weren't new data to be shared.

Code Availability	Yes (indicate where provided:	n/a
For all newly generated code and software essential for replicating the main findings of the study:		
State whether the code or software is available.	Data Sharing statement	
If code is publicly available, provide accession number in repository, or DOI or URL.		n/a Emails could be sent to the address below to obtain the shared code: aijunqin2022@126.com.

Reporting

Adherence to community standards	Yes (indicate where provided: section/paragraph)	n/a

MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.	MDAR and TRIPOD reporting checklists were provided.	
State if relevant guidelines (eg., ICMJE, MIBBI, ARRIVE) have been followed, and whether a checklist (eg., CONSORT, PRISMA, ARRIVE) is provided with the manuscript.	ICMJE guidelines were followed, as the journal follows ICMJE recommendations for publication.	

Article information: https://dx.doi.org/10.21037/atm-22-6659