Appendix 2. RECORD and TRIPOD checklists

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found		RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Title Title NA
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction		
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction		
Methods					
Study Design	4	Present key elements of study design early in the paper	Study Design		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Study Design		

Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Inclusion and exclusion criteria, Data collection Data collection NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.		RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Data collection
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Data collection, Supplementary data		
Bias	9	Describe any efforts to address	Statistical analysis,		

		potential sources of bias	discussion		
Study size	10	Explain how the study size was arrived at	NA		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Statistical analysis, Supplementary data		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Statistical analysis		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Study design and setting, Other information Supplementary

					data
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Study design and setting
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram		RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.*, study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Inclusion and exclusion criteria, Figure 1
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)			Table 1
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report			Table 2

		numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			Table 3, Table 4, Supplementary data
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses			Supplementary data (notably Table S9 and S10)
Discussion					
Key results	18	Summarise key results with reference to study objectives	Key results		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Limitations and strengths
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Interpretation and generalisability		

Generalisability	21	Discuss the generalisability (external validity) of the study results	Interpretation and generalisability		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Funding		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Accessibility of protocol, raw data, and programming code

TRIPOD Checklist: Prediction Model Validation

Section/Topic	ı	Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	3
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	3
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	4
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	4,5,8
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	Describe eligibility criteria for participants.	4,5
	5c	Give details of treatments received, if relevant.	NA
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	5,6,7
	6b	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	Sup.
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	Explain how the study size was arrived at.	NA
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	6, Sup.
Statistical analysis methods	10c	For validation, describe how the predictions were calculated.	6,7, Sup
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	6,7
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	Provide details on how risk groups were created, if done.	NA
Development vs. validation	12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Sup.
Results			
Participants	13a	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.	Fig. 1
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	Table 1.
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Table 1. and Sup
Model performance	16	Report performance measures (with CIs) for the prediction model.	Table 3 and 4, Sup
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	10,11
Interpretation	19a	For validation, discuss the results with reference to performance in the development data, and any other validation data.	11,12
	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	11,12
Implications	20	Discuss the potential clinical use of the model and implications for future research.	10,11, 12
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	13
Funding	22	Give the source of funding and the role of the funders for the present study.	13