Pub	Med
#1	"stomach neoplasms"[MeSH Terms] OR ("stomach"[Title/Abstract] AND
	"neoplasms"[Title/Abstract]) OR "stomach neoplasms"[Title/Abstract] OR
	("gastric"[Title/Abstract] AND "cancer"[Title/Abstract]) OR "gastric"
	cancer"[Title/Abstract] OR ("stomach neoplasms"[MeSH Terms] OR
	("stomach"[Title/Abstract] AND "neoplasms"[Title/Abstract]) OR "stomach
	neoplasms"[Title/Abstract] OR ("stomach"[Title/Abstract] AND
	"cancer"[Title/Abstract]) OR "stomach cancer"[Title/Abstract]) OR
	(("gastrics"[Title/Abstract] OR "stomach"[MeSH Terms] OR
	"stomach"[Title/Abstract] OR "gastric"[Title/Abstract]) AND
	("carcinoma"[MeSH Terms] OR "carcinoma"[Title/Abstract] OR
	"carcinomas"[Title/Abstract] OR "carcinoma s"[Title/Abstract])) OR
	(("stomach"[MeSH Terms] OR "stomach"[Title/Abstract] OR
	"stomachs"[Title/Abstract] OR "stomach s"[Title/Abstract] OR
	"stomachal"[Title/Abstract] OR "stomaches"[Title/Abstract]) AND
	("carcinoma" [MeSH Terms] OR "carcinoma" [Title/Abstract] OR
	"carcinomas"[Title/Abstract] OR "carcinoma s"[Title/Abstract])) OR ("stomach
	neoplasms"[MeSH Terms] OR ("stomach"[Title/Abstract] AND
	"neoplasms"[Title/Abstract] OR "stomach neoplasms"[Title/Abstract] OR
	("gastric"[Title/Abstract] AND "neoplasm"[Title/Abstract]) OR "gastric neoplasm"[Title/Abstract]) OR ("stomach neoplasms"[MeSH Terms] OR
	("stomach"[Title/Abstract] AND "neoplasms"[Title/Abstract]) OR "stomach neoplasms"[Title/Abstract] OR ("stomach"[Title/Abstract] AND
	• · · · · · · · · · · · · · · · · · ·
	"neoplasm"[Title/Abstract]) OR "stomach neoplasm"[Title/Abstract]) OR
	("stomach neoplasms"[MeSH Terms] OR ("stomach"[Title/Abstract] AND
	"neoplasms"[Title/Abstract]) OR "stomach neoplasms"[Title/Abstract] OR ("stomach"[Title/Abstract] AND "tumor"[Title/Abstract]) OR "stomach
	tumor"[Title/Abstract]) OR ("stomach neoplasms"[MeSH Terms] OR
	("stomach"[Title/Abstract] AND "neoplasms"[Title/Abstract]) OR "stomach
	neoplasms"[Title/Abstract] OR ("gastric"[Title/Abstract] AND "tumor"[Title/Abstract])
#2	(("artificial intelligence"[MeSH Terms] OR ("artificial"[Title/Abstract] AND
π∠	"intelligence"[Title/Abstract]) OR "artificial intelligence"[Title/Abstract]) OR
	("deep learning"[MeSH Terms] OR ("deep"[Title/Abstract] AND
	"learning"[Title/Abstract]) OR "deep learning"[Title/Abstract])) OR ("machine
	learning [MeSH Terms] OR ("machine" [Title / Abstract] AND
	"learning"[Title/Abstract]) OR "machine learning"[Title/Abstract])
#3	"mortality"[MeSH Subheading] OR "mortality"[Title/Abstract] OR
113	"survival" [Title/Abstract] OR "survival" [MeSH Terms] OR
	"survivability"[Title/Abstract] OR "survivable"[Title/Abstract] OR
	"survivals" [Title/Abstract] OR "survive" [Title/Abstract] OR
	"survived" [Title / Abstract] OR survives [Title / Abstract] OR "survives" [Title / Abstract] OR
	"survived [Title/Abstract] OK Survives [Title/Abstract] OK Survives [Title/Abstract]
	#1 AND #2 AND #3 Results: N=442
Wah	of Science
#1	(((((((((TS=(survival)) OR TS=(mortality)) AND TS=(gastric cancer)) OR
π1	
	TS=(stomach cancer)) OR TS=(gastric neoplasm)) OR TS=(stomach neoplasm)) OR TS=(gastric tumor)) OR TS=(gastric carcinoma))
	OR TS=(gastric tumor)) OR TS=(stomach tumor)) OR TS=(gastric carcinoma)) OR TS=(stomach carcinoma)
	ON 10-(Stomach Carchoma)

#2	((TS=(machine	learning))	OR	TS=(deep	learning))	OR	TS=(artificial
	intelligence)						
#3	(TS=(survival))	OR TS=(mor	tality)				
	#1 AND #2 AN	D#3 Resul	ts: N=	935			

Supplementary Table S2: CHARMS Checklist.

Domain	S A Rahman (2021)	Reported on page #
SOURCE OF DATA	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location, number of centers, setting, inclusion and exclusion criteria)	2
PARTICIPANTS	Participant description	2,4
	Details of treatments received, if relevant	2,4
	Study dates	2,4
	Definition and method for measurement of outcome	2
	Was the same outcome definition (and method for measurement) used in all patients?	2
OUTCOME(S) TO	Type of outcome (e.g., single or combined endpoints)	2
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
	Number and type of predictors (e.g., demographics, patient history, physical examination, additional testing, disease characteristics)	2-3
CANDIDATE	Definition and method for measurement of candidate predictors	2-3
PREDICTORS	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation)	2-3
(OR INDEX TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
	Handling of predictors in the modelling (e.g., continuous, linear, non-linear transformations or categorised)	2-3
CANADI E CIZE	Number of participants and number of outcomes/events	3
SAMPLE SIZE	Number of outcomes/events in relation to the number of candidate predictors (Events Per Variable)	3
	Number of participants with any missing value (include predictors and outcomes)	2
MISSING DATA	Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning techniques)	3-4
	Modelling assumptions satisfied	3-4
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all candidate predictors, pre-selection based on unadjusted association with the outcome)	3-4
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model approach, backward or forward selection) and criteria used (e.g., p-value, Akaike Information Criterion)	3-4
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform shrinkage, penalized estimation)	3-4
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination (C-statistic, D-statistic, log-rank) measures with confidence intervals	3-4
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification improvement) and whether a-priori cut points were used	3-4
MODEL EVALUATION	Method used for testing model performance: development dataset only (random split of data, resampling methods e.g. bootstrap or cross-validation, none) or separate external validation (e.g. temporal, geographical, different setting, different investigators)	4
	In case of poor validation, whether model was adjusted or updated (e.g., intercept recalibrated, predictor effects adjusted, or new predictors added)	4

	Final and other multivariable models (e.g., basic, extended, simplified) presented, including predictor weights or regression coefficients, intercept, baseline survival, model performance measures (with standard errors or confidence intervals)	4-6
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram, score chart, predictions for specific risk subgroups with performance	4-6
	Comparison of the distribution of predictors (including missing data) for development and validation datasets	4-6
INTERPRETATION AND DISCUSSION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus exploratory, i.e., more research needed)	6-8
AND DISCUSSION	Comparison with other studies, discussion of generalizability, strengths and limitations.	6-8

Domain	Tao Chen (2019)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
PARTICIPANTS	number of centers, setting, inclusion and exclusion criteria)	
TARTICITATION	Participant description	2,4
	Details of treatments received, if relevant	2,4
	Study dates	2,4
	Definition and method for measurement of outcome	2
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	2
	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)? Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
		2-3
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2-3
PREDICTORS	additional testing, disease characteristics) Definition and method for measurement of candidate predictors	2-3
	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
(OR INDEX	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
TESTS)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2-3
	transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	2
	Number of outcomes/events in relation to the number of candidate predictors (Events	2
A 41CCINIC DATA	Number of participants with any missing value (include predictors and outcomes)	2
MISSING DATA	Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	2-3
	Modelling assumptions satisfied	2-3
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	2-3
DEVELOPMENT	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPIVILINI	Method for selection of predictors during multivariable modelling (e.g., full model	2-3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	2-3
	shrinkage, penalized estimation)	_
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	3
PERFORMANCE	Classification massures (o.g. consitiuity, anasificity, and distinguished and a second second	2
	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	3
	improvement) and whether a-priori cut points were used	3
MODEL	Method used for testing model performance: development dataset only (random split of	3
EVALUATION	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
	validation (e.g. temporal, geographical, different setting, different investigators) In case of poor validation, whether model was adjusted or updated (e.g., intercept	3
]
	recalibrated, predictor effects adjusted, or new predictors added)	<u> </u>

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	4		
	including predictor weights or regression coefficients, intercept, baseline survival, model			
DECLUTO	performance measures (with standard errors or confidence intervals)			
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	4		
	score chart, predictions for specific risk subgroups with performance			
	Comparison of the distribution of predictors (including missing data) for development and	4		
	validation datasets			
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	5-6		
AND DISCUSSION	AND DISCUSSION exploratory, i.e., more research needed)			
	Comparison with other studies, discussion of generalizability, strengths and limitations.	5-6		

Domain	Mengxin Tian (2024)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2-3
PARTICIPANTS	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	2-3,7
	Details of treatments received, if relevant	2-3,7
	Study dates	2-3
	Definition and method for measurement of outcome	3
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	3
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	3
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	3
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	3
	Time of outcome occurrence or summary of duration of follow-up	3
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	3
PREDICTORS	additional testing, disease characteristics)	_
PREDICTORS	Definition and method for measurement of candidate predictors	3
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	3
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	3
12313)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	3
	transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	2-3
	Number of outcomes/events in relation to the number of candidate predictors (Events	3
MISSING DATA	Number of participants with any missing value (include predictors and outcomes)	3
	Number of participants with missing data for each predictor	3
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	3
	Modelling method (e.g., logistic, survival, neural network, or machine learning	3-4
	Modelling assumptions satisfied	3-4
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	3-4
DEVELOPMENT	candidate predictors, pre-selection based on unadjusted association with the outcome)	2.4
	Method for selection of predictors during multivariable modelling (e.g., full model	3-4
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	2.4
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	3-4
	shrinkage, penalized estimation)	
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	4
PERFORMANCE	Classification management (a.g. consitiuity, angelisity, and distingualization and an electrication	1
	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	4
	improvement) and whether a-priori cut points were used	1
MODEL	Method used for testing model performance: development dataset only (random split of	4
EVALUATION	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
	validation (e.g. temporal, geographical, different setting, different investigators)	4
	In case of poor validation, whether model was adjusted or updated (e.g., intercept	4
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	5-11
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	5-11
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	5-11
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	11-13
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	11-13

Domain	Warid Islam (2024)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location, number of centers, setting, inclusion and exclusion criteria)	2
PARTICIPANTS	Participant description	2
	Details of treatments received, if relevant	2
	Study dates	2
	Definition and method for measurement of outcome	2
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2
OUTCOME(S) TO	Type of outcome (e.g., single or combined endpoints)	2
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination, additional testing, disease—characteristics)	2-3
PREDICTORS	Definition and method for measurement of candidate predictors	2-3
(OD INIDEV	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
(OR INDEX	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
TESTS)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear transformations or categorised)	2-3
SAMPLE SIZE	Number of participants and number of outcomes/events	2
37 (1411 EE 312E	Number of outcomes/events in relation to the number of candidate predictors (Events	2
	Number of participants with any missing value (include predictors and outcomes)	2
MISSING DATA	Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	2-3
	Modelling assumptions satisfied	2-3
140051	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	2-3
MODEL	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	2-3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	2-3
	shrinkage, penalized estimation)	
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	3
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	3
	improvement) and whether a-priori cut points were used	
	Method used for testing model performance: development dataset only (random split of	3
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	
EVALUATION	In case of poor validation, whether model was adjusted or updated (e.g., intercept	3
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	4
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	4
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	4
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	5-6
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	5-6

Hongcai Chen (2024)	Reported
	on page #
Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
number of centers, setting, inclusion and exclusion criteria)	
Participant description	2
Details of treatments received, if relevant	2
Study dates	2
Definition and method for measurement of outcome	2
Was the same outcome definition (and method for measurement) used in all patients?	2
	2
	2
	2
Time of outcome occurrence or summary of duration of follow-up	2
Number and type of predictors (e.g., demographics, patient history, physical examination,	2
additional testing, disease characteristics)	_
	2
	2
·	2
	2
	2
	2
	2
	2
	2
	2
	2
	2
	2
	2
	2,4-5
Cambration (cambration plot, cambration slope, nosmer-terneshow test) and discrimination	2,4-5
Classification measures (e.g. sensitivity specificity predictive values net reclassification	2,4-5
	_,. 5
	2
validation (e.g. temporal, geographical, different setting, different investigators)	
In case of poor validation, whether model was adjusted or updated (e.g., intercept	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location, number of centers, setting, inclusion and exclusion criteria) Participant description Details of treatments received, if relevant Study dates Definition and method for measurement of outcome Was the same outcome definition (and method for measurement) used in all patients? Type of outcome (e.g., single or combined endpoints) Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)? Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)? Time of outcome occurrence or summary of duration of follow-up Number and type of predictors (e.g., demographics, patient history, physical examination,

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	3-5
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	3-5
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	3-5
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	6-9
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	6-9

Domain	Naoki Kuwayama (2023)	Reported
Domain		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
DADTICIDANITO	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	2
	Details of treatments received, if relevant	2
	Study dates	2
	Definition and method for measurement of outcome	2
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2
	Type of outcome (e.g., single or combined endpoints)	2
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2
	additional testing, disease characteristics)	
PREDICTORS	Definition and method for measurement of candidate predictors	2
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2
•	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2
TESTS)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2
	transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	2
	Number of outcomes/events in relation to the number of candidate predictors (Events	2
MISSING DATA	Number of participants with any missing value (include predictors and outcomes)	2
WIISSING DATA	Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	2-3
	Modelling assumptions satisfied	2-3
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	2-3
	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	2-3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	2-3
	shrinkage, penalized estimation)	
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	4-5
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	4-5
	improvement) and whether a-priori cut points were used	
MODEL	Method used for testing model performance: development dataset only (random split of	2-3
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	
Z V / LO / (TIO)	In case of poor validation, whether model was adjusted or updated (e.g., intercept	2-3
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	3-5
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTE	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	3-5
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	3-5
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	5-9
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	5-9

Domain	Junjie Zeng (2024)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
DARTICIDANTO	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	2,4
	Details of treatments received, if relevant	2,4
	Study dates	2,4
	Definition and method for measurement of outcome	2
011700145(6) 70	Was the same outcome definition (and method for measurement) used in all patients?	2
OUTCOME(S) TO	Type of outcome (e.g., single or combined endpoints)	2
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2
PREDICTORS	additional testing, disease characteristics)	_
PREDICTORS	Definition and method for measurement of candidate predictors	2
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2
12313)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2
SAMPLE SIZE	transformations or categorised) Number of participants and number of outcomes/events	2
SAIVIPLE SIZE	Number of outcomes/events in relation to the number of candidate predictors (Events	2
	Number of participants with any missing value (include predictors and outcomes)	2
MISSING DATA	Number of participants with any missing value (meduce predictors and outcomes) Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	2-3
	Modelling assumptions satisfied	2-3
	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	2-3
MODEL	candidate predictors, pre-selection based on unadjusted association with the outcome)	2 3
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	2-3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	2 3
	Shrinkage of predictor weights or regression coefficients (e.g., p-value, Akaike	2-3
	shrinkage, penalized estimation)	2 3
	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	5-7
MODEL	Cambration (Cambration plot, Cambration Slope, Hostiler-Lemeshow test) and Discrimination	J-7
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	5-7
	improvement) and whether a-priori cut points were used	
	Method used for testing model performance: development dataset only (random split of	3
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION:	validation (e.g. temporal, geographical, different setting, different investigators)	
EVALUATION	In case of poor validation, whether model was adjusted or updated (e.g., intercept	3
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	4-7
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTE	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	4-7
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	4-7
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	7-10
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	7-10

Domain	Mengjie Wu (2024)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
PARTICIPANTS	number of centers, setting, inclusion and exclusion criteria)	
7,4416.174.415	Participant description	2,5
	Details of treatments received, if relevant	2,5
	Study dates	2,5
	Definition and method for measurement of outcome	2-3
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2-3
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	2-3
	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2-3
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2-3
	Time of outcome occurrence or summary of duration of follow-up	2-3
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2-3
PREDICTORS	additional testing, disease characteristics) Definition and method for measurement of candidate predictors	2-3
	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
(OR INDEX	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
TESTS)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2-3
	transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	2-3
	Number of outcomes/events in relation to the number of candidate predictors (Events	2-3
	Number of participants with any missing value (include predictors and outcomes)	2-3
MISSING DATA	Number of participants with missing data for each predictor	2-3
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2-3
	Modelling method (e.g., logistic, survival, neural network, or machine learning	3-4
	Modelling assumptions satisfied	3-4
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	3-4
	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	3-4
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	3-4
	shrinkage, penalized estimation)	
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	10-12
PERFORMANCE		
TENTONIVIANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	10-12
	improvement) and whether a-priori cut points were used	
MODEL	Method used for testing model performance: development dataset only (random split of	3-4
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	2.4
	In case of poor validation, whether model was adjusted or updated (e.g., intercept	3-4

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	4-5
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTE	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	4-5
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	4-5
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	5-8
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	5-8

Domain	Xunjun Li (2022)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
PARTICIPANTS	number of centers, setting, inclusion and exclusion criteria)	
TARTICITATION	Participant description	2-3
	Details of treatments received, if relevant	2-3
	Study dates	2-3
	Definition and method for measurement of outcome	2
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	2
DE I NEDICIED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2
PREDICTORS	additional testing, disease characteristics)	2
TREDICTORS	Definition and method for measurement of candidate predictors	2
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2
•	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2
CANADI E CIZE	transformations or categorised)	2
SAMPLE SIZE	Number of participants and number of outcomes/events Number of outcomes/events in relation to the number of candidate predictors (Events	2
MISSING DATA	Number of participants with any missing value (include predictors and outcomes) Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning Modelling assumptions satisfied	2-3 2-3
	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	2-3
MODEL		2-3
DEVELOPMENT	candidate predictors, pre-selection based on unadjusted association with the outcome) Method for selection of predictors during multivariable modelling (e.g., full model	2-3
		2-3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	2-3
		2-3
	shrinkage, penalized estimation) Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	4-5
MODEL	Cambration (cambration plot, cambration slope, hosiner-terneshow test) and discrimination	4-3
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	4-5
	improvement) and whether a-priori cut points were used	. 3
	Method used for testing model performance: development dataset only (random split of	2
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators) In case of poor validation, whether model was adjusted or updated (e.g., intercept	2
	recalibrated, predictor effects adjusted, or new predictors added)	-

	end and other configuration and the configuration of the configuration o	2.5
	Final and other multivariable models (e.g., basic, extended, simplified) presented,	2-5
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTE	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	2-5
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	2-5
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	5-7
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	5-7

Domain	Rocío Aznar-Gimeno (2024)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	3
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	3
DADTICIDANITO	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	3,7-8
	Details of treatments received, if relevant	3,7-8
	Study dates	3,7-8
	Definition and method for measurement of outcome	3
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	3
	Type of outcome (e.g., single or combined endpoints)	3
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	3
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	3
	Time of outcome occurrence or summary of duration of follow-up	3
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	3-4
	additional testing, disease characteristics)	
PREDICTORS	Definition and method for measurement of candidate predictors	3-4
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	3-4
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	3-4
12313)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	3-4
	transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	3
	Number of outcomes/events in relation to the number of candidate predictors (Events	3
MISSING DATA	Number of participants with any missing value (include predictors and outcomes)	3
WIISSING DATA	Number of participants with missing data for each predictor	3
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	3
	Modelling method (e.g., logistic, survival, neural network, or machine learning	4-7
	Modelling assumptions satisfied	4-7
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	4-7
DEVELOPMENT	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPIVIENT	Method for selection of predictors during multivariable modelling (e.g., full model	4-7
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	4-7
	shrinkage, penalized estimation)	
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	11-13
PERFORMANCE		
I LINI ONIVIAINCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	11-13
	improvement) and whether a-priori cut points were used	
MODEL	Method used for testing model performance: development dataset only (random split of	4-7
IVIODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	
	In case of poor validation, whether model was adjusted or updated (e.g., intercept	4-7
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	7-17
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTE	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	7-17
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	7-17
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	17-21
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	17-21

Domain	Yuming Jiang (2022)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
PARTICIPANTS	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	2,5
	Details of treatments received, if relevant	2,5
	Study dates	2,5
	Definition and method for measurement of outcome	2
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	2
DE FREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2-3
PREDICTORS	additional testing, disease characteristics)	2.2
TREDICTORS	Definition and method for measurement of candidate predictors	2-3
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
,	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2-3
CANADI E CIZE	transformations or categorised) Number of participants and number of outcomes/events	3
SAMPLE SIZE	Number of outcomes/events in relation to the number of candidate predictors (Events	3
	Number of participants with any missing value (include predictors and outcomes)	2
MISSING DATA	Number of participants with any missing value (meduce predictors and outcomes)	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	3-4
	Modelling assumptions satisfied	3-4
	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	3-4
MODEL	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	3-4
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	3-4
	shrinkage, penalized estimation)	
	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	3-4
MODEL		
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	3-4
	improvement) and whether a-priori cut points were used	
	Method used for testing model performance: development dataset only (random split of	4
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	
LVALOATION	In case of poor validation, whether model was adjusted or updated (e.g., intercept	4
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	4-6
	including predictor weights or regression coefficients, intercept, baseline survival, model	. 0
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	4-6
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	4-6
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	6-8
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	6-8

Domain	Fan Li (2024)	Reported
Domain	Tail Li (2024)	on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
DA DELCIDANIEC	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	2
	Details of treatments received, if relevant	2
	Study dates	2
	Definition and method for measurement of outcome	2
OUTCOME(C) TO	Was the same outcome definition (and method for measurement) used in all patients?	2
OUTCOME(S) TO	Type of outcome (e.g., single or combined endpoints)	2
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2
	Time of outcome occurrence or summary of duration of follow-up	2
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2-3
PREDICTORS	additional testing, disease characteristics)	
FILDICIONS	Definition and method for measurement of candidate predictors	2-3
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
123137	Handling of predictors in the modelling (e.g., continuous, linear, non-linear transformations or categorised)	2-3
SAMPLE SIZE	Number of participants and number of outcomes/events	2
SAIVIPLE SIZE	Number of outcomes/events in relation to the number of candidate predictors (Events	2
	Number of participants with any missing value (include predictors and outcomes)	2
MISSING DATA	Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	2-3
	Modelling assumptions satisfied	2-3
	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	2-3
MODEL	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	2-3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	2-3
	shrinkage, penalized estimation)	
	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	3-4
MODEL	Canada (Canada) prod, canada (Canada)	
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	3-4
	improvement) and whether a-priori cut points were used	
	Method used for testing model performance: development dataset only (random split of	2-3
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	
	In case of poor validation, whether model was adjusted or updated (e.g., intercept	2-3
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	5-8
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	5-8
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	5-8
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	8-9
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	8-9

Domain	Tianbao Liao (2024)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2-3
PARTICIPANTS	number of centers, setting, inclusion and exclusion criteria)	
FARTICIFANTS	Participant description	2-3,5
	Details of treatments received, if relevant	2-3,5
	Study dates	2-3,5
	Definition and method for measurement of outcome	3
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	3
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	3
DETREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	3
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	3
	Time of outcome occurrence or summary of duration of follow-up	3
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2-3
PREDICTORS	additional testing, disease characteristics)	2.2
TREDICTORS	Definition and method for measurement of candidate predictors	2-3
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
,	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2-3
SAMPLE SIZE	transformations or categorised) Number of participants and number of outcomes/events	2
SAIVIPLE SIZE	Number of outcomes/events in relation to the number of candidate predictors (Events	2
	Number of participants with any missing value (include predictors and outcomes)	2
MISSING DATA	Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	3
	Modelling assumptions satisfied	3
	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	3
MODEL	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	3
	shrinkage, penalized estimation)	
MODE	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	6-7
MODEL		
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	6-7
	improvement) and whether a-priori cut points were used	
	Method used for testing model performance: development dataset only (random split of	3
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	
LVALUATION	In case of poor validation, whether model was adjusted or updated (e.g., intercept	3

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	4-6
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	4-6
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	4-6
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	6-10
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	6-10

Domain	Ting Wei (2022)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2-3
DARTICIDANITO	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	2-3,6
	Details of treatments received, if relevant	2-3,6
	Study dates	2-3,6
	Definition and method for measurement of outcome	2-3
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2-3
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	2-3
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2-3
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2-3
	Time of outcome occurrence or summary of duration of follow-up	2-3
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2-3
PREDICTORS	additional testing, disease characteristics)	
PREDICTORS	Definition and method for measurement of candidate predictors	2-3
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
12313)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2-3
	transformations or categorised)	_
SAMPLE SIZE	Number of participants and number of outcomes/events	2
	Number of outcomes/events in relation to the number of candidate predictors (Events	2
MISSING DATA	Number of participants with any missing value (include predictors and outcomes)	2
	Number of participants with missing data for each predictor	2
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2
	Modelling method (e.g., logistic, survival, neural network, or machine learning	3-4
	Modelling assumptions satisfied	3-4
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	3-4
DEVELOPMENT	candidate predictors, pre-selection based on unadjusted association with the outcome)	2.4
	Method for selection of predictors during multivariable modelling (e.g., full model	3-4
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	2.4
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	3-4
	shrinkage, penalized estimation)	_
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	7
PERFORMANCE	Classification management (a.g. consitiuity, angelisity, and distingualization and an electrication	7
	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	'
	improvement) and whether a-priori cut points were used	2
MODEL	Method used for testing model performance: development dataset only (random split of	3
	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	3
	In case of poor validation, whether model was adjusted or updated (e.g., intercept	3
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	6-9
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	6-9
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	6-9
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	9-11
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	9-11

Domain	Mohammad Reza Afrash (2023)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	3
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	3-4
DARTICIDANTO	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	3-4
	Details of treatments received, if relevant	3-4
	Study dates	3-4
	Definition and method for measurement of outcome	4
a=aa=(a) =a	Was the same outcome definition (and method for measurement) used in all patients?	4
OUTCOME(S) TO	Type of outcome (e.g., single or combined endpoints)	4
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	4
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	4
	Time of outcome occurrence or summary of duration of follow-up	4
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	4
CANDIDATE	additional testing, disease characteristics)	
PREDICTORS	Definition and method for measurement of candidate predictors	4
(OD INIDEV	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	4
(OR INDEX	Were predictors assessed blinded for outcome, and for each other (if relevant)?	4
TESTS)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	4
	transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	3
	Number of outcomes/events in relation to the number of candidate predictors (Events	3
A ALCCINIC DATA	Number of participants with any missing value (include predictors and outcomes)	3
MISSING DATA	Number of participants with missing data for each predictor	3
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	3
	Modelling method (e.g., logistic, survival, neural network, or machine learning	5-6
	Modelling assumptions satisfied	5-6
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	5-6
MODEL	candidate predictors, pre-selection based on unadjusted association with the outcome)	
DEVELOPMENT	Method for selection of predictors during multivariable modelling (e.g., full model	5-6
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	5-6
	shrinkage, penalized estimation)	
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	9-12
MODEL		
PERFORMANCE	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	9-12
	improvement) and whether a-priori cut points were used	
	Method used for testing model performance: development dataset only (random split of	6
MODEL	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION:	validation (e.g. temporal, geographical, different setting, different investigators)	
EVALUATION	In case of poor validation, whether model was adjusted or updated (e.g., intercept	6
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	6-12
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTO	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	6-12
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	6-12
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	12-14
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	12-14

Domain	Junjie Zeng (2023)	Reported
		on page #
SOURCE OF	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	2
	Participant eligibility and recruitment method (e.g., consecutive participants, location,	2
DADTICIDANITO	number of centers, setting, inclusion and exclusion criteria)	
PARTICIPANTS	Participant description	2,5
	Details of treatments received, if relevant	2,5
	Study dates	2,5
	Definition and method for measurement of outcome	2-3
OUTCOME(S) TO	Was the same outcome definition (and method for measurement) used in all patients?	2-3
BE PREDICTED	Type of outcome (e.g., single or combined endpoints)	2-3
BE PREDICTED	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	2-3
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	2-3
	Time of outcome occurrence or summary of duration of follow-up	2-3
CANDIDATE	Number and type of predictors (e.g., demographics, patient history, physical examination,	2-3
PREDICTORS	additional testing, disease characteristics)	
PREDICTORS	Definition and method for measurement of candidate predictors	2-3
(OR INDEX	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment	2-3
TESTS)	Were predictors assessed blinded for outcome, and for each other (if relevant)?	2-3
12313)	Handling of predictors in the modelling (e.g., continuous, linear, non-linear	2-3
	transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	2-3
	Number of outcomes/events in relation to the number of candidate predictors (Events	2-3
MISSING DATA	Number of participants with any missing value (include predictors and outcomes)	2-3
	Number of participants with missing data for each predictor	2-3
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	2-3
	Modelling method (e.g., logistic, survival, neural network, or machine learning	3
	Modelling assumptions satisfied	3
MODEL	Method for selection of predictors for inclusion in multivariable modelling (e.g., all	3
DEVELOPMENT	candidate predictors, pre-selection based on unadjusted association with the outcome)	2
5272201 WEW	Method for selection of predictors during multivariable modelling (e.g., full model	3
	approach, backward or forward selection) and criteria used (e.g., p-value, Akaike	2
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform	3
	shrinkage, penalized estimation)	0.11
MODEL	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination	9-11
PERFORMANCE		0.11
	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification	9-11
	improvement) and whether a-priori cut points were used	2
MODEL	Method used for testing model performance: development dataset only (random split of	3
	data, resampling methods e.g. bootstrap or cross-validation, none) or separate external	
EVALUATION	validation (e.g. temporal, geographical, different setting, different investigators)	3
	In case of poor validation, whether model was adjusted or updated (e.g., intercept	3
	recalibrated, predictor effects adjusted, or new predictors added)	

	Final and other multivariable models (e.g., basic, extended, simplified) presented,	4-7
	including predictor weights or regression coefficients, intercept, baseline survival, model	
DECLUTE	performance measures (with standard errors or confidence intervals)	
RESULTS	Any alternative presentation of the final prediction models, e.g., sum score, nomogram,	4-7
	score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and	4-7
	validation datasets	
INTERPRETATION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus	7-11
AND DISCUSSION	exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	7-11