

Appendix 1 Searching strategies, inclusion and exclusion criteria, and quality assessment criteria

Searching strategies

Pubmed

- #1 "Mortality"[Mesh]
- #2 "Death"[Mesh]
- #3 ((Mortalit*[Title/Abstract]) OR (Fatalit*[Title/Abstract])) OR (Death*[Title/Abstract])
- #4 (("Mortality"[Mesh]) OR ("Death"[Mesh])) OR (((Mortalit*[Title/Abstract]) OR (Fatalit*[Title/Abstract])) OR (Death*[Title/Abstract]))
- #5 (NEWS2[Title/Abstract]) OR (National Early Warning Score 2[Title/Abstract])
- #6 (National Early Warning Score 2) OR (NEWS2)
- #7 (((("Mortality"[Mesh]) OR ("Death"[Mesh])) OR (((Mortalit*[Title/Abstract]) OR (Fatalit*[Title/Abstract])) OR (Death*[Title/Abstract])))) AND ((National Early Warning Score 2) OR (NEWS2))

Embase

- #1 'mortality'/exp
- #2 'death'/exp
- #3 'fatality'/exp
- #4 mortalit*:ab,ti OR fatalit*:ab,ti OR death*:ab,ti
- #5 #1 OR #2 OR #3 OR #4
- #6 'national early warning score 2'/exp
- #7 'national early warning score 2' OR news2
- #8 #6 OR #7
- #9 #5 AND #8

Cochrane Library

- #1 MeSH descriptor: [Mortality] explode all trees
- #2 MeSH descriptor: [Death] explode all trees
- #3 (Mortalit*):ti,ab,kw OR (Fatalit*):ti,ab,kw OR (Death*):ti,ab,kw
- #4 (National Early Warning Score 2) OR (NEWS2)
- #5 #1 OR #2 OR #3
- #6 #4 AND #5

Web of Science

- #1 National Early Warning Score 2 (Topic) or NEWS2

(Topic)

- #2 TS=(Mortalit*) OR TS=(Fatalit*) OR TS=(Death*)
- #3 #2 AND #1

CNKI

- #1 ((旧版主题=中英文扩展(NEWS2)+中英文扩展('National Early Warning Score 2') 或者 keyword=NEWS2+'National Early Warning Score 2' 或者 title=NEWS2+'National Early Warning Score 2' 或者 abstract=NEWS2+'National Early Warning Score 2') 并且 (旧版主题=死亡率 或者 keyword=中英文扩展(死亡率) 或者 title=中英文扩展(死亡率) 或者 abstract=中英文扩展(死亡率))) (模糊匹配), 专辑导航: 全部; 数据库: 文献跨库检索

WanFang Data

- #1 全部:(NEWS2 or 'National Early Warning Score 2') and 全部:(死亡率)

Vip Database

- #1 任意字段=NEWS2 OR 'National Early Warning Score 2' AND 任意字段=死亡率

SinoMed

- #1 "NEWS2"[全部字段:智能] AND "死亡率"[全部字段:智能]

Quality Assessment Criteria

Two authors used the Prediction model Risk Of Bias Assessment Tool (PROBAST) independently to assess the risk of bias of included trials. The PROBAST consists of assessment of four key domains to judge the quality of studies: participants, predictors, outcome, analysis. The answer to each item was "+", "-", or "?" ("+" indicates low risk of bias; "-" indicates high risk of bias; and "?" indicates unclear risk of bias). If a study was judged as "low" on all domains relating to bias, then it was assigned an overall judgment of "low risk of bias" or "low concern regarding applicability", and had high quality. If a study was judged "high" in one or more domains, then it may have been judged as "at risk of bias" or "concerns regarding applicability". Disagreements were resolved by the third author.

Appendix 2 Details of the quality assessment

Authors	Risk of Bias																							Applicability				
	1. Participants			2. Predictors			3. Outcome						4. Analysis							1. Participants	2. Predictors	3. Outcome	(1, 2, 3) Applicability					
	1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?	1.2 Were all inclusions and exclusions of participants appropriate?	overall	2.1 Were predictors defined and assessed in a similar way for all participants?	2.2 Were predictor assessments made without knowledge of outcome data?	2.3 Are all predictors available at the time the model is intended to be used?	overall	3.1 Was the outcome determined appropriately?	3.2 Was a pre-specified or standard outcome definition used?	3.3 Were predictors excluded from the outcome definition?	3.4 Was the outcome defined and determined in a similar way for all participants?	3.5 Was the outcome determined without knowledge of predictor information?	3.6 Was the time interval between predictor assessment and outcome determination appropriate?	overall	4.1 Were there a reasonable number of participants with the outcome??	4.2 Were continuous and categorical predictors handled appropriately?	4.3 Were all enrolled participants included in the analysis?	4.4 Were participants with missing data handled appropriately?	★ 4.5 Was selection of predictors based on univariable analysis avoided?	4.6 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?	4.7 Were relevant model performance measures evaluated appropriately?	★ 4.8 Were model overfitting and optimism in model performance accounted for?		★ 4.9 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?	overall	(1, 2, 3, 4)Risk of Bias	1. Concern that the included participants and setting do not match the review question.	2. Concern that the definition, assessment or timing of predictors in the model do not match the review question.
Medina-Lozano <i>et al.</i> 2020 (18)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2020 (19)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Marincowitz <i>et al.</i> 2022 (23)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Hu <i>et al.</i> 2022 (24)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Guarino <i>et al.</i> 2022 (25)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Chikhalkar <i>et al.</i> 2022 (26)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Villanueva Rabano <i>et al.</i> 2021 (27)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Thomas <i>et al.</i> 2021 (28)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Sivayoham <i>et al.</i> 2021 (29)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Richardson <i>et al.</i> 2021 (30)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Reardon <i>et al.</i> 2021 (31)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Prasad <i>et al.</i> 2021 (32)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Osawa <i>et al.</i> 2021 (33)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Masson <i>et al.</i> 2021 (34)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2021 (35)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2021 (36)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+

Appendix 2 (continued)

Authors	Risk of Bias																							Applicability				
	1. Participants			2. Predictors			3. Outcome						4. Analysis							1. Participants	2. Predictors	3. Outcome	(1, 2, 3) Applicability					
	1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?	1.2 Were all inclusions and exclusions of participants appropriate?	overall	2.1 Were predictors defined and assessed in a similar way for all participants?	2.2 Were predictor assessments made without knowledge of outcome data?	2.3 Are all predictors available at the time the model is intended to be used?	overall	3.1 Was the outcome determined appropriately?	3.2 Was a pre-specified or standard outcome definition used?	3.3 Were predictors excluded from the outcome definition?	3.4 Was the outcome defined and determined in a similar way for all participants?	3.5 Was the outcome determined without knowledge of predictor information?	3.6 Was the time interval between predictor assessment and outcome determination appropriate?	overall	4.1 Were there a reasonable number of participants with the outcome??	4.2 Were continuous and categorical predictors handled appropriately?	4.3 Were all enrolled participants included in the analysis?	4.4 Were participants with missing data handled appropriately?	★ 4.5 Was selection of predictors based on univariable analysis avoided?	4.6 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?	4.7 Were relevant model performance measures evaluated appropriately?	★ 4.8 Were model overfitting and optimism in model performance accounted for?		★ 4.9 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?	overall	(1, 2, 3, 4)Risk of Bias	1. Concern that the included participants and setting do not match the review question.	2. Concern that the definition, assessment or timing of predictors in the model do not match the review question.
Lopez-Izquierdo <i>et al.</i> 2021 (37)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Durantez-Fernandez <i>et al.</i> 2022 (38)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Durantez-Fernandez <i>et al.</i> 2021 (39)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Clair <i>et al.</i> 2021 (40)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Mellhammar <i>et al.</i> 2020 (41)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2020 (42)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2020 (43)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Magnusson <i>et al.</i> 2020 (44)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Covino <i>et al.</i> 2020 (45)	N	Y	-	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	-	Y	Y	Y	+
Mellhammar <i>et al.</i> 2019 (46)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2019 (47)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2019 (48)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2019 (49)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+
Martin-Rodriguez <i>et al.</i> 2019 (50)	Y	Y	+	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	+	Y	Y	Y	Y	Y	Y	Y	Y	+	+	Y	Y	Y	+

★ : This question is limited to model development studies; “Y” indicates positive of the question; “N” indicates negative of the question ; “+” indicates low risk of bias; “-” indicates high risk of bias.