BMJ Open

Prognostic factors and prediction models for acute aortic dissection: a systematic review

Yan Ren [ID] ,[1] Shiyao Huang,[1] Qianrui Li,[1,2] Chunrong Liu,[1] Ling Li,[1] Jing Tan,[1] Kang Zou,[1] Xin Sun[1]

[1]Chinese Evidence-based Medicine Center and National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Chengdu, Sichuan, China
[2]Department of Nuclear Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Correspondence to
Professor Xin Sun;
sunxin@wchscu.cn

ABSTRACT

Objective Our study aimed to systematically review the methodological characteristics of studies that identified prognostic factors or developed or validated models for predicting mortalities among patients with acute aortic dissection (AAD), which would inform future work.

Design/setting A methodological review of published studies.

Methods We searched PubMed and EMBASE from inception to June 2020 for studies about prognostic factors or prediction models on mortality among patients with AAD. Two reviewers independently collected the information about methodological characteristics. We also documented the information about the performance of the prognostic factors or prediction models.

Results Thirty-two studies were included, of which 18 evaluated the performance of prognostic factors, and 14 developed or validated prediction models. Of the 32 studies, 23 (72%) were single-centre studies, 22 (69%) used data from electronic medical records, 19 (59%) chose retrospective cohort study design, 26 (81%) did not report missing predictor data and 5 (16%) that reported missing predictor data used complete-case analysis. Among the 14 prediction model studies, only 3 (21%) had the event per variable over 20, and only 5 (36%) reported both discrimination and calibration statistics. Among model development studies, 3 (27%) did not report statistical methods, 3 (27%) exclusively used statistical significance threshold for selecting predictors and 7 (64%) did not report the methods for handling continuous predictors. Most prediction models were considered at high risk of bias. The performance of prognostic factors showed varying discrimination (AUC 0.58 to 0.95), and the performance of prediction models also varied substantially (AUC 0.49 to 0.91). Only six studies reported calibration statistic.

Conclusions The methods used for prognostic studies on mortality among patients with AAD—including prediction models or prognostic factor studies—were suboptimal, and the model performance highly varied. Substantial efforts are warranted to improve the use of the methods in this population.

Strengths and limitations of this study

► This systematic review study is the first to identify methodological gaps and assess the performance of the prognostic factors or prediction models among all studies addressing individual prognostic factors or developing or validating prediction models on mortality among patients with acute aortic dissection (AAD).

► This review designed a comprehensive questionnaire that included items from both Prediction model Risk Of Bias ASsessment Tool (PROBAST) and CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) checklists and assessed methodological gaps among all studies.

► This review is important that the methodological quality of models designed to support medical decision for patients with AAD, substantial efforts are warranted to strengthen the use of rigorous methods for the accuracy and reliability of the performance in the future research.

► The small number of prediction models limits the recommendation in clinical practice, combining international registry of acute aortic dissection (IRAD) score and C-reactive protein model showed better discrimination than IRAD score, future studies may consider updating IRAD model by including other relevant biomarkers, which may further improve prognostic performance.

► Our review about the methodological characteristics was primarily based on reporting, which might be cases that the researchers had considered the methodological issues but did not clearly report.

INTRODUCTION

Acute aortic dissection (AAD) is a life-threatening cardiovascular disease with high mortality, characterised with acute onset and rapid progression. The mortality of untreated AAD was approximately 1%–2% per hour early following the onset of symptoms, and the overall in-hospital mortality was approximately 27%.[1 2] Treatment options for AAD include medical intervention, surgery or endovascular repair, the selection of which mainly depends on complications and prognosis of patients.[3] Better understanding of the disease prognosis, ideally predicting the risk of a serious outcome, is highly desirable

for medical decision-making and patient communication, among which mortality has the highest priority.

Several published systematic reviews assessed the association of inflammatory biomarkers (eg, C-reactive protein (CRP)) and marker of cardiac injury (ie, troponin) with increased mortality in patients with AAD.[4–6] A few studies also developed or validated prediction models for mortality in AAD,[7–9] in which a combination of biomarkers, demographic and clinical characteristics were included.[8 10–14] As a result, they have received increasing use in clinical practice.

However, limited efforts have been made to systematically examine the performance of the prognostic factors or prediction models. In particular, a comprehensive assessment is strongly needed to investigate whether the published studies—either individual prognostic factor studies or prediction models—meet the desirable methodological rigours for clinical use, since suboptimal methods can compromise the accuracy and reliability of the risk estimation. This is particularly the case for AAD, a disease condition, whereby predictability of an adverse outcome has paramount importance. Therefore, we conducted a systematic review study to identify methodological gaps among all studies addressing individual prognostic factors or developing or validating prediction models on mortality among patients with AAD.

METHODS

We conducted this systematic review according to a prespecified protocol, which was not published.

Eligibility criteria

We developed the eligibility criteria under the Population, Index prognostic factor, Comparator prognostic factors, Outcome, Timing and Setting (PICOTS) guidance.[15] A study was eligible for inclusion if it included patients diagnosed with AAD; and aimed to identify or assess any prognostic factors for mortality, or develop or validate a prognostic model for mortality in patients with AAD. We excluded a study if it was prediction model for AAD diagnosis only; or the report was a review, comment, letter or editorial, case report, protocol or conference abstract.

Predictors measured at any time point in the course of AAD were eligible. No restriction on study setting was applied; patients with AAD who visited any healthcare facilities were eligible. We defined a prognostic prediction model as a multivariable model, predicting risk of specific outcomes occurring in future by selected predictors.[16]

Literature search and screening

We searched PubMed and EMBASE from inception to June 2020 for relevant reports published in English language. We conducted the search using the Medical Subject Headings (MeSH) terms and free texts to identify reports about AAD, including 'aortic dissecting aneurysm', 'aortic aneurysm', 'aortic dissection*' and 'aortic dissecting hematoma'. We applied a validate search strategy for searching prediction

models, which proved to have high sensitivity and specificity.[17] The full search strategy is presented as online supplemental appendix A. Two investigators (YR and SH) independently screened all searched reports, and resolved any disagreements through discussion with a third investigator (CL). We also manually searched for additional articles from the reference lists of all selected articles.

Data extraction

We collected the following general information from each eligible study, including first author, year of publication, study aim, region of study, type of aortic dissection, age and sex ratio. We carefully collected information about performance of identified prognostic factors or prediction models, including their names and results about discrimination, calibration, sensitivity and specificity. Discrimination and calibration are the two key measures for evaluating the predictive performance of the prognostic factors or prediction models.[18]

In order to examine the methods used among these prognoses studies, a team of methods-trained, experienced methodologists expertise with prognostic studies and prediction models convened to develop a questionnaire through a consensus process. They first consulted items from the published statements and tools (eg, PROBAST, CHARMS checklist) about prognoses studies,[19 20] and brainstormed for additional items. Subsequently, they discussed the identified items about their relevance for methods, and dropped items that were deemed irrelevant. Finally, they achieved consensus about the items through group discussion and agreement.

Generally, this questionnaire consists of five domains: (1) **study design** (number of centres, sample size, number of events, data sources, epidemiological design); (2) **participants** (definition and selection of participants); (3) **predictors** (definition and measurement of predictors); (4) **outcome** (definition and measurement of outcomes) and (5) **analysis** (were all enroled participants included in the analysis, the number of events per variable (EPV), statistical method for selecting and handling predictors, missing data, model structure used in the study and relevant model performance measures evaluated for addressing prognostic factors or prediction models). The questionnaire was presented as in online supplemental appendix B.

Additionally, we used a risk of bias assessment tool adapted from the PROBAST tool to assess the risk of bias for prediction modelling studies.[15 20] The detailed tool and assessment criteria are presented in online supplemental appendix C.

Statistical analysis

Categorical variables were expressed as the number of frequencies and proportion. For quantitative variables, data were summarised by mean and SD or median with IQR according to normality tests.
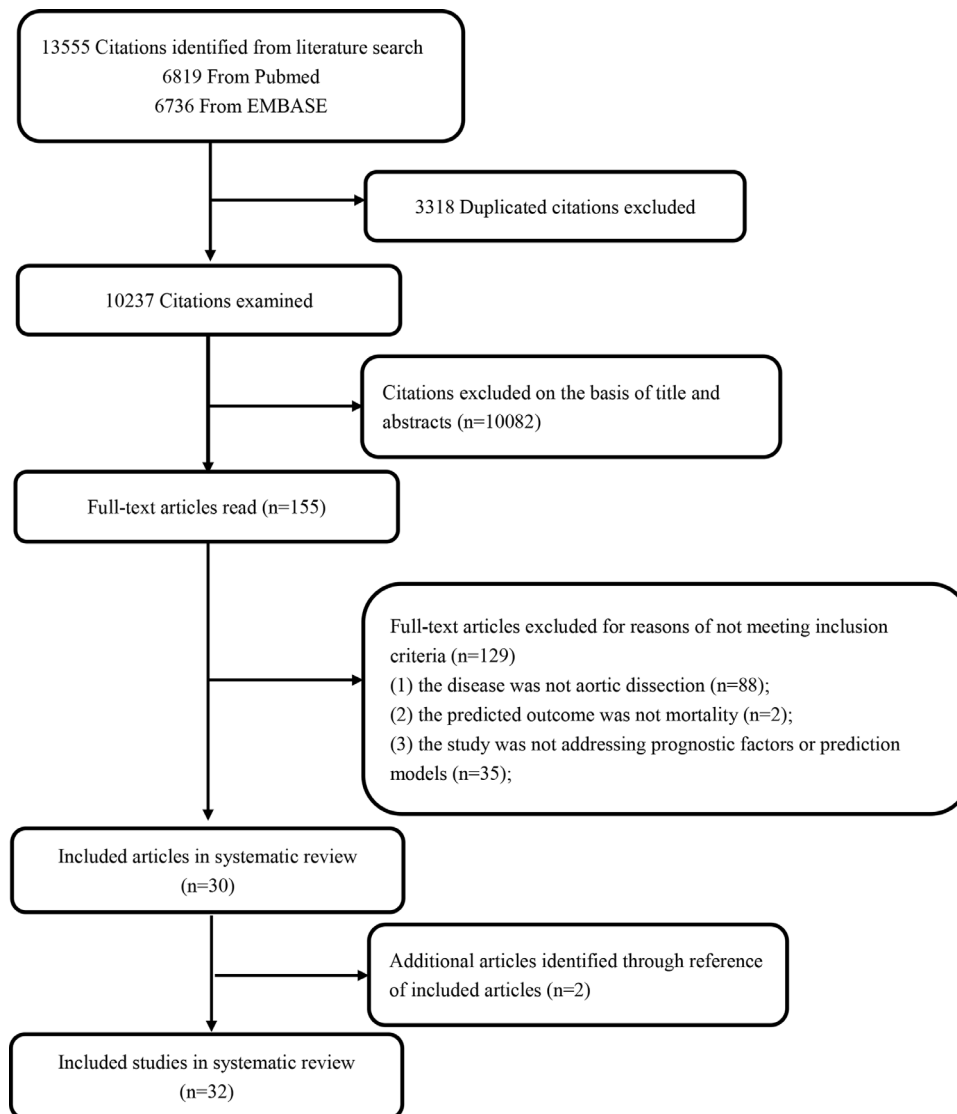
Figure 1 Flow chart of study selection.

RESULTS

In total, 13 555 records were identified, among which 155 were selected for full-text screening, and 32 studies were eligible and included in the final analysis (figure 1).

General characteristics of included studies

The 32 eligible studies were published between 2002 and 2019 (online supplemental appendix table 1). Five (15%) were multinational studies, and 21 (66%) were conducted in the USA, China and Europe. The dissection types of patients with AAD were mostly Type-A (n=21, 66%), followed by a mixture of Type-A and Type-B (n=8, 25%). In-hospital mortality was the most frequently used outcome (n=24, 75%, table 1).

Eighteen (56%) studies aimed to evaluate the performance of prognostic factors. The most commonly investigated prognostic factors were D-dimer (DD, n=8), neutrophil lymphocyte ratio (NLR, n=4) and CRP (n=3). Fourteen (44%) studies aimed to develop or validate a prediction model, of which nine developed a new prediction model without any validation, two developed a new

prediction model with internal validation and three conducted external validation with or without updating a prediction model (table 1).

Model performance

The performance of prognostic factors showed poor to strong discrimination (AUC 0.58 to 0.95). The AUC of single prognostic factor ranged from 0.58 to 0.92, and the one for combined prognostic factors ranged from 0.77 to 0.95 (DD and CRP: 0.95; NT-proBNP and aortic diameter: 0.83; Tenascin-C (TNC) and DD: 0.95; TNC and CRP: 0.91; cystatin C and high-sensitivity C-reactive protein: 0.88; UA, DD and age: 0.77, table 2).

The developed or validated models from 11 studies showed poor to strong discrimination (AUC 0.49 to 0.91), only 6 reported calibrations, and of which 5 reported good calibrations (p>0.05). Rampoldi et al developed a prediction model and reported moderate discrimination (AUC 0.76). But through external validation, scoring systems developed by Rampoldi et al showed poor discrimination (30-day mortality: AUC 0.56, operative mortality: AUC 0.62). Mehta

Table 1 General characteristics about design and conduct of studies

Characteristics	Number (%)
Study region	
One country	27 (84.4)
China	14 (43.8)
USA	3 (9.4)
Europe	4 (12.5)
Other	5 (15.6)
Multinational	5 (15.6)
Multicentre study	
Yes	9 (28.1)
No	23 (71.9)
The most commonly reported prognostic biomarkers (n=18)	
D-dimer	8 (44.4)
NLR	4 (22.2)
CRP	3 (16.7)
Study purpose	
Identification or assessment of prognostic factors	18 (56.2)
Development or validation of a prediction models	14 (43.8)
Develop a model without validation	9 (28.1)
Develop a model internal validation	2 (6.3)
External validation	3 (9.4)
Dissection type	
A	21 (65.6)
B	3 (9.4)
A/B	8 (25.0)
Outcome (some studies have more than one outcome, such as in-hospital mortality and 1 year mortality)	
In-hospital mortality	24 (75.0)
Operative mortality	2 (6.25)
30-Day mortality	4 (12.5)
Long-term mortality (included 1 year mortality)	5 (15.6)

CRP, C-reactive protein; NLR, neutrophil lymphocyte ratio.

et al (p value for the Hosmer-Lemeshow (H-L) test=0.75) developed a prediction model using International Registry of Acute Aortic Dissection (IRAD) from multinational data and reported good calibration. Through external validation, IRAD score showed moderate discrimination (AUC 0.74), addition of CRP to IRAD score notably improved discrimination (AUC 0.89, table 2).

Methodological characteristics
Among the 32 studies, most were single-centre studies (n=23, 72%). The sample size varied from 35 to 1034 (median 165, IQR, 103–348), and the median number of events was 35 (23–72). Thirteen (41%) studies used prospective cohort study design, and the rest 19 (59%) used retrospective cohort study design, 22 (69%) used data from electronic medical records, 5 (16%) from cohort studies and 5 (16%) from registries (table 3).

Thirty-one (97%) studies clearly described inclusion and exclusion criteria for participants. The criteria used to define and to measure predictors in the study population were consistent among all included studies. The criteria for outcome definition and measurement was consistent in all but one study[13] (table 3).

Twenty-two (69%) studies included all enroled participants in the analysis. In the handling of missing data, 30 (94%) studies reported no missing outcome data, 26 (81%) did not report missing predictor data and 5 (16%) reported that there were some predictors with missing data, and used complete-case analysis to handle missing predictors (table 3).

In 18 prognostic factor studies, 9 (50%) had the EPV more than 20, 8 (44%) between 10 and 20 and 1 (6%) less than 10; 15 (83%) reported discrimination, sensitivity and specificity, other 3 (17%) only reported discrimination, or sensitivity and specificity; and 11 (61%) chose logistic regression model for the analysis, 5 (28%) used cox regression, 2 (11%) only used Receiver Operating Characteristic (ROC) analysis (table 3).

In the 14 prediction model studies, only 3 (21%) had the EPV more than 20, 8 (57%) between 10 and 20 and 3 (21%) less than 10; 10 (71%) chose logistic regression model for the analysis, other four studies used cox regression, support vector machines, neural networks and ROC analysis, respectively. The performance measures were poorly reported: only five (36%) reported both discrimination and calibration statistics. Eleven (64%) studies reported discrimination, measured as AUC of the receiver operated curve, and six (43%) reported calibration, measured as p value for the H-L test. For developing a prediction model, three (27%) did not report any statistical methods and three (27%) simply used statistical significance for selecting predictors; seven (64%) did not report how to handle continuous predictors, four (36%) reported continuous predictor was transformed into categories (table 3).

Risk of bias assessment
The risk of bias for 14 prediction models in the domains of participants, predictors and outcome was low for most studies, while the risk of bias in the domain of sample size and missing data and statistical analysis was generally high (table 4). Studies rated high and unclear risk of bias in the domains of sample size and missing data, due to low number of outcomes per variable (EPV <10), or lack of information about the method on handling missing data. The main reasons for studies rated high and unclear risk of bias in the domains of statistical analysis are as follows: the predictors are selected on the basis of univariable analysis prior to multivariable modelling, lack of information on whether continuous predictors are examined for non-linearity and how categorical predictor groups are

Table 2 Reported discrimination and calibration of prognostic factors or prediction models for AAD

Study ID	Dissection type	Predictor	Outcome	AUC (95% CI)	P value of Hosmer-Lemeshow test	Sensitivity (%)	Specificity (%)
Prognostic factors							
Liu et al (2018)[27]	A	Fibrinogen	In-hospital mortality	0.686 (0.585–0.787)		71.90	60.40
Zindovic et al (2018)[28]	A	Preoperative lactic acid levels	In-hospital mortality	0.684		56.00	72.00
			1 year mortality	0.673		48.00	74.00
		Postoperative lactic acid levels	In-hospital mortality	0.582			
			1 year mortality	0.498			
Oz et al (2017)[29]	A	NLR	In-hospital mortality	0.919 (0.832–1.00)		86.00	91.00
Feng et al (2017)[30]	A	Serum cystatin C	Long-term mortality (followed up for 909 days)	0.772 (0.692–0.839)		78.53	69.23
		hs-CRP		0.640 (0.574–0.739)		86.72	46.51
		Cystatin C, hs-CRP		0.883 (0.826–0.935)		97.44	65.92
Li et al (2016)[11]	A	hs-TnT	Long-term mortality (followed up for 3.5 years)	0.719 (0.621–0.803)		70.80	76.40
		hs-CRP		0.700 (0.599–0.789)		48.90	94.30
		D-dimer		0.818 (0.724–0.891)		86.10	71.40
Karakoyun et al (2015)[31]	A	NLR	In-hospital mortality	0.829 (0.674–0.984)		77.00	74.00
Wen et al (2019)[14]	A/B	NT-proBNP	In-hospital mortality	0.799 (0.707–0.891)		55.20	95.70
		Aortic diameter		0.724 (0.607–0.841)		58.60	88.20
		NT-proBNP and aortic diameter		0.832 (0.735–0.929)		79.30	84.90
Liu et al (2018)[32]	A/B	BUN	In-hospital mortality	0.785 (0.662–0.909)		78.90	72.20
Bennett et al (2017)[33]	A	Serum lactic acid level	In-hospital mortality	0.88		85.00	77.00
			1 year mortality	0.81		67.00	84.00
Lafçi et al (2014)[34]	A/B	NLR	In-hospital mortality	0.634 (0.516–0.753)		70.00	53.00
Wen et al (2013)[13]	A/B	D-dimer	In-hospital mortality	0.917 (0.85–0.96)		90.30	75.90
		CRP		0.822 (0.74–0.89)		100.00	54.20
		D-dimer + CRP		0.948 (0.89–0.98)		81.90	96.80
Guo et al (2019)[10]	A/B	TNC	In-hospital mortality	0.884 (0.809–0.937)		83.87	83.33
		TNC +D-dimer		0.946 (0.885–0.980)		90.30	88.46
		D-dimer		0.787 (0.698–0.859)		87.19	64.10
		CRP		0.758 (0.667–0.835)		90.32	55.13
		TNC + CRP		0.909 (0.839–0.956)		90.32	74.92
Ohlmann et al (2006)[12]	A/B	D-dimer	In-hospital mortality	0.650 (0.584–0.716)			

Continued

Table 2 Continued

Study ID	Dissection type	Predictor	Outcome	AUC (95% CI)	P value of Hosmer-Lemeshow test	Sensitivity (%)	Specificity (%)
Zhang et al (2016)[35]	A	WBC	In-hospital mortality			84.60	65.90
		SBP				65.90	69.20
		NT-proBNP				80.80	51.20
		D-dimer				84.60	70.70
Li et al (2019)[36]	B	PLR	In-hospital mortality	0.711 (0.580–0.840)		63.00	88.00
Zhang et al (2020)[37]	A	UA	In-hospital mortality	0.678 (0.579–0.777)		65.00	67.10
		D-dimer		0.689 (0.589–0.790)		44.70	88.80
		age		0.616 (0.507–0.724)		37.50	90.40
		UA, D-dimer, age		0.771			
Bedel et al (2019)[38]	A	NLR	In-hospital mortality	0.746 (0.623–0.870)		70.60	76.80
		PLR		0.750 (0.638–0.882)		76.50	78.10
Gong et al (2019)[39]	A	Postoperative TnI	30-Day mortality	0.711			
		Postoperative Mb		0.699			
		Preoperative CK-MB		0.694			
		Postoperative CK-MB		0.678			
		Preoperative Creatinine		0.668			
		Preoperative Mb		0.644			
		Preoperative D-Dimer		0.621			
		Preoperative TnI		0.618			
Prediction models							
Develop a model without validation							
Zhang et al (2015)[40]	A/B	Hypotension, syncope, ischaemic complications, renal dysfunction, type A, neutrophil percentage ≥80%, surgery	In-hospital mortality	0.650	0.160		
Tolenaar et al (2014)[8]	B	Female, age, hypotension/shock, periaortic haematoma, aortic diameter ≥5.5 cm, mesenteric ischaemia, acute renal failure, limb ischaemia	In-hospital mortality		p=0.314		
Mehta et al (2002)[7]	A	Age, female, abrupt onset pain, abnormal ECG, any pulse deficit, kidney failure, hypotension/shock/tamponade	In-hospital mortality	0.740	p=0.750		
Ghoreishi et al (2018)[41]	A	Lactic acid, creatinine, liver malperfusion	Operative mortality	0.750			
Centofanti et al (2006)[42]	A	Age, coma, acute renal failure, shock and redo operation	30-Day mortality	Only reported the expected mortality and observed mortality			

Continued

Table 2 Continued

Study ID	Dissection type	Predictor	Outcome	AUC (95% CI)	P value of Hosmer-Lemeshow test	Sensitivity (%)	Specificity (%)
Santini et al (2007)[43]	A	Age, cardiac tamponade, hypotension, acute myocardial ischaemia, mesenteric ischaemia, acute renal failure, neurologic injury	In-hospital mortality	0.763 (0.802–0.723)		55.60	82.90
Rampoldi et al (2007)[44]	A	Age >70, history of aortic valve replacement,hypotension (systolic blood pressure <100mm Hg) or shock at presentation,migrating chest pain, preoperative cardiac tamponade,any pulse deficit,ECG with findings of myocardial ischaemia or infarction	In-hospital mortality	0.760	p=0.230		
	A	Age >70, history of aortic valve replacement,hypotension (systolic blood pressure <100mm Hg) or shock at presentation,migrating chest pain, preoperative cardiac tamponade, any pulse deficit, intraoperative hypotension, right ventricle dysfunction at surgery, a necessity to perform a coronary artery bypass graft		0.810	p=0.380		
Leontyev et al (2016)[45]	A	Age, critical preoperative state, malperfusion syndrome, coronary artery disease	In-hospital mortality	0.767 (0.715–0.819)	p=0.60		
Zhang et al (2019)[46]	B	Hypotension, Ischaemic complications, renal dysfunction, neutrophil percentage	In-hospital mortality			86 (risk score ≥4)	78 (risk score ≥4)
Develop a model with internal validation							
Macrina et al (2010)[47]	A	Immediate postoperative chronic renal failure, circulatory arrest time, the type of surgery on ascending aorta plus hemi-arch, extracorporeal circulation time and the presence of Marfan habitus	Long-term mortality (564±48 days)	Support vector machines:0.821, neural networks: 0.870			

Continued

Table 2 Continued

Study ID	Dissection type	Predictor	Outcome	AUC (95% CI)	P value of Hosmer-Lemeshow test	Sensitivity (%)	Specificity (%)
Macrina et al (2009)[48]	A	Immediate postoperative presence of dialysis in continuous, renal complications, chronic renal failure, coded operative brain protection (anterograde better than retrograde perfusion), preoperative neurological symptoms, age, previous cardiac surgery, the length of extracorporeal circulation, the operative presence of haemopericardium and postoperative enterological complications	30-Day mortality	First centre: multiple logistic regression 0.879 (0.807–0. 932)			
		Immediate postoperative presence of chronic renal failure, coded operative brain protection (anterograde better than retrograde perfusion), postoperative presence of dialysis in continuous, preoperative neurological symptoms, postoperative renal complications, the length of extracorporeal circulation, age, the operative presence of haemopericardium, preoperative presence of intubation, postoperative limb ischaemia and enterological complications and the year of surgery		Second centre: multiple logistic regression 0.857 (CI: 0.785 to 0.911) Second centre: neural networks 0.905 (0.838–0.951)			
External validation							
Ge et al (2013)[49]	A/B	EuroSCORE II	In-hospital mortality	0.490 (0.390–0.590)	p<0.001		
Yu et al (2016)[50]	A	Scoring systems developed by Rampoldi et al	Operative mortality	0.62			
			30-Day mortality	0.56			
		Scoring systems developed by Centofanti et al	Operative mortality	0.66			
			30-Day mortality	0.58			
		Age	Operative mortality	0.67			
Vrsalovic et al (2015)[9]	A	CRP	In-hospital mortality	0.790 (0.784–0.796)		83.00	80.00
		IRAD score		0.740 (0.733–0.747)			
		IRAD score + CRP		0.890 (0.886–0.894)			

Rampoldi et al scoring system was calculated for each patient as −3.20 + (0.68 × age >70) + (1.44 × history of aortic valve replacement) + (1.17 × hypotension or shock at presentation) + (0.88 × migrating chest pain) + (0.97 × preoperative cardiac tamponade) + (0.56 × any pulse deficit) + (0.57 × ECG with findings of myocardial ischaemia or infarction).

Centofanti et al scoring system was calculated for each patient as: −2.986 + (0.771 × shock) + (0.595 × reoperation) + (1.162 × coma) + (0.778 × acute renal failure) + (0.023 × age).

AAD, acute aortic dissection; BUN, blood urea nitrogen; CK-MB, creatine kinase MB isoenzyme; CRP, C-reactive protein; hs-CRP, high-sensitivity C-reactive protein; hs-TnT, high-sensitivity cardiac troponin T; EuroSCORE II, European System for Cardiac Operative Risk Evaluation; Mb, myoglobin; NLR, neutrophil lymphocyte ratio; NT-proBNP, N-terminal pro-brain natriuretic peptide; PLR, Platelet count to lymphocyte count ratio; IRAD score, international registry of acute aortic dissection score; TNC, Tenascin-C; UA, Uric Acid.

Table 3 Methodological characteristics of included studies	
Characteristics	**Number (%) or median (IQR)**
Sample size (n)	165 (103–348)
Death events (n)	35 (23–72)
Multicentre study	
Yes	9 (28.1)
No	23 (71.9)
Epidemiological design	
Prospective cohort	13 (40.6)
Retrospective cohort	19 (59.4)
Data sources	
Cohort study	5 (15.6)
EMR data	22 (68.8)
Registry	5 (15.6)
Did the study clearly describe inclusion/ exclusion criteria for participants?	
Yes	31 (96.9)
No	1 (3.1)
Consistent definition/diagnostic criteria of predictors used in all participants	
Yes	32 (100.0)
No	0 (0)
Consistent measurement of predictors used in all participants	
Yes	32 (100.0)
No	0 (0)
Consistent definition/diagnostic criteria of outcomes used in all participants	
Yes	31 (96.9)
No	1 (3.1)
Consistent measurement of outcomes used in all participants	
Yes	31 (96.9)
No	1 (3.1)
Were all enroled participants included in the analysis?	
Yes	22 (68.8)
No	10 (31.2)
Was missing outcome data reported, and the methods for handling missing outcome	
Yes, complete-case analysis	1 (3.1)
No	30 (93.8)
Not reported	1 (3.1)
Was any missing predictor data reported, and the methods for handling missing predictor	
Yes, complete-case analysis	5 (15.6)
No	1 (3.1)
	Continued

Table 3 Continued	
Characteristics	**Number (%) or median (IQR)**
Not reported	26 (81.3)
Prognostic factors (n=18) prediction models	
Number of outcomes/events in relation to the number of predictors for assessing prognostic factors (EPVs)	
<10	1 (5.6)
10–20	8 (44.4)
≥20	9 (50.0)
Model structure used in the study	
Logistic regression	11 (61.1)
Cox regression	5 (27.8)
ROC analyses (not report regression)	2 (11.1)
Relevant model performance measures evaluated for addressing prognostic factors	
AUC	2 (11.1)
AUC, sensitivity, specificity	15 (83.3)
Sensitivity, specificity	1 (5.6)
Prediction models (n=14)	
Number of outcomes/events in relation to the number of predictors in multivariable analysis (EPVs)	
<10	3 (21.4)
10–20	8 (57.1)
≥20	3 (21.4)
Model structure used in the study	
Logistic regression	10 (71.4)
Cox regression	1 (7.1)
ROC analyses (not report regression)	1 (7.1)
Logistic regression and support vector machines	1 (7.1)
Logistic regression and neural networks	1 (7.1)
Relevant model performance measures evaluated for addressing prediction models	
AUC, p value of Hosmer-Lemeshow test	5 (35.7)
AUC	4 (28.6)
AUC, sensitivity, specificity	2 (14.3)
P value of Hosmer-Lemeshow test	1 (7.1)
Expected and observed	1 (7.1)
Sensitivity, specificity	1 (7.1)
Develop prediction models (n=11)	
Statistical method for selecting predictors during addressing prediction models	
Univariate analysis of predictors by P value	3 (27.3)
Univariate analysis of predictors by p value and other specific predictors	3 (27.3)
	Continued

Table 3	Continued
Characteristics	**Number (%) or median (IQR)**
Stepwise selection	2 (18.1)
Not reported	3 (27.3)
Handling the predictors for addressing prediction models	
Continuous predictor was transformed into categories	4 (36.4)
Not reported	7 (63.6)

.EMRs, electronic medical records; EPV, events per variable.

defined and either calibration or discrimination are not reported.

DISCUSSION
Summary study findings
In this systematic review, we identified 32 studies addressing prognostic factors or prediction models for mortality among patients with AAD . As noticed in this review, the performance of prognostic factors or prediction models was most commonly evaluated by the AUC and H-L test. Most assessment of prognostic factors demonstrated moderate discrimination. The factors using combined TNC and DD, or combined DD and CRP showed strong discrimination (AUC 0.95). The prediction models showed poor to strong discrimination (AUC 0.49 to 0.91). The prediction model European System for Cardiac Operative Risk Evaluation (EuroSCORE II) showed poor discriminative ability (AUC 0.49) and poor calibration (p value for the H-L test <0.001). One explanation may be that Euro-SCORE II is a risk model which allows the calculation of the risk of death after a heart surgery, and is not related to prognosis of patients with AAD, because not all patients with aortic dissection undergo surgical treatment, and some of them undergo endovascular treatment. Mehta *et al's*[7] model showed better discrimination (0.74) than the EuroSCORE II. Meanwhile, Mehta *et al* used IRAD from multinational data reported good calibration. Through external validation, IRAD score showed moderate discrimination (AUC 0.74), addition of CRP to IRAD score notably improved discrimination (AUC 0.89). Hence, the prediction model for mortality in AAD should consider including biomarkers as predictors to improve discrimination.

In this systematic review, we found that most studies had small number of sample sizes and events, were derived from a single-centre study and a relatively large proportion of studies chose to use retrospective data. Most studies did not describe information on missing data nor accounted for appropriate statistical methods for handling missing data.

For developing or validating prediction models, we found that most were considered at high risk of bias; the number of EPV in most studies was relatively small, which result in prediction performance of models being possibly biased[21 22]; most studies did not evaluate both discrimination and calibration. Almost all studies reported discriminative ability of prediction models, while only six studies reported calibration. For developing prediction models, we found that some studies based on statistical significance for selecting variable may lead to suboptimal models; most studies did not report how to handle the continuous variable, and linear assumption may be inappropriate.

Table 4 Risk of bias of included prediction model studies

Study ID	Participants	predictors	Outcome	Sample size and missing data	Statistical analysis
Zhang *et al* (2015)[40]	L	L	L	H	H
Tolenaar *et al* (2014)[8]	L	L	L	H	H
Mehta *et al* (2002)[7]	L	L	L	U	U
Ghoreishi *et al* (2018)[41]	L	L	H	U	H
Centofanti *et al* (2006)[42]	L	L	L	U	H
Santini *et al* (2007)[43]	L	L	L	U	H
Rampoldi *et al* (2007)[44]	L	L	L	L	H
Leontyev *et al* (2016)[45]	L	L	L	U	H
Zhang *et al* (2019)[46]	L	L	L	H	H
Macrina *et al* (2010)[47]	L	L	L	H	H
Macrina *et al* (2009)[48]	L	L	L	H	H
Ge *et al* (2013)[49]	H	H	L	H	H
Yu *et al* (2016)[50]	L	L	L	H	H
Vrsalovic *et al* (2015)[9]	L	L	L	H	H

L, low risk; H, high risk; U, unclear risk.

Implications for future study

Although some studies showed good discrimination and calibration, our findings highlighted important methodological limitations among those studies. Then it is possible that the result is not accurate and reliable. So in the future, studies about prognostic factors or prediction models for mortality in AAD should enrol large patient population from multicentre setting, meanwhile consider cohort designs, the imputation of missing data. Multiple imputation techniques to deal with missing data are important when evaluating model performance. Excluding cases with missing data may lead to biased results.[23]

Studies about prediction models for mortality in AAD should consider appropriate methods for selecting variable and handling the continuous variable, and evaluating both discrimination and calibration. The number of participants and events should be planned, and the number of EPV should be at least 10. If the number of events is low relative to the number of predictors, penalised regression may be better than the standard regression. Stability selection and subsampling have demonstrated to yield more stable models based on a consistent selection of variables, so they should be used in future studies for prediction model.[24] Discrimination should not be reported in isolation because a poorly calibrated model can present the same discriminative capacity as a perfectly calibrated one.[25] Reporting both discrimination and calibration is highly recommended for evaluating performance measures. Validating the prediction models should be considered, as both model development and validation are essential processes for establishing a useful prediction model.[26]

A prediction model most suitable for clinical practice should include a relatively small number of variables, be easily interpreted and have good statistical performance. Apart from the well-established IRAD model, our review found that the combined IRAD score and CRP model used less variables and showed better discrimination than IRAD score alone. These characteristics may warrant daily practice of the combine model. Moreover, future studies may consider updating IRAD model by including other relevant biomarkers, which may further improve prognostic performance in clinical practice.

Strengths and limitations

To our knowledge, no systematic review looking at the methodology characteristics and performance of prognostic factors or predictive models for mortality in AAD has been published. Whether these existing prognostic factors or prediction models may be used to guide or improve clinical practice remains underexplored. Should we seek better prognostic factors or prediction models? Should we continue using and validating these prognostic factors or prediction models? There is consensus on this issue among commentators. We should seek better prognostic factors or prediction models. Substantial efforts are warranted to strengthen the use of rigorous methods for the accuracy and reliability of the performance in the future research.

A limitation of the present study is that our review about the methodological characteristics was primarily based on reporting. There might be cases that the researchers had considered the methodological issues but did not clearly report. This situation also emphasised the importance of complete reporting.

CONCLUSIONS

In conclusion, DD, NLR and CRP predictors were the most commonly used biomarkers, the performance of prognostic factors showed a poor to strong discrimination, the prediction models varied substantially, only six studies reported the calibration, and of which five reported good calibration. Meanwhile, many of these prognostic factors or predictive models are weak methodologically, several important issues are needed to consider for strengthening for predicting mortality in AAD, such as the sample size, the methods for handling missing data, appropriate statistical analysis methods and reporting both calibration and discrimination for prediction models. Substantial efforts are warranted to improve the use of the methods for better care of this population.

ORCID iD
Yan Ren http://orcid.org/0000-0002-1937-5019

REFERENCES

1. Nienaber CA, Eagle KA. Aortic dissection: new frontiers in diagnosis and management: Part I: from etiology to diagnostic strategies. *Circulation* 2003;108:628–35.
2. Hagan PG, Nienaber CA, Isselbacher EM, *et al*. The International registry of acute aortic dissection (IRAD): new insights into an old disease. *JAMA* 2000;283:897–903.
3. Nienaber CA, Clough RE. Management of acute aortic dissection. *The Lancet* 2015;385:800–11.
4. Vrsalovic M. Prognostic effect of cardiac troponin elevation in acute aortic dissection: a meta-analysis. *Int J Cardiol* 2016;214:277–8.
5. Hsieh WC, Henry BM, Hsieh CC, *et al*. Prognostic role of admission C-reactive protein level as a predictor of in-hospital mortality in type-A acute aortic dissection: a meta-analysis. *Vasc Endovascular Surg* 2019;53:547–57.
6. Vrsalović M, Vrsalović Presečki A. Admission C-reactive protein and outcomes in acute aortic dissection: a systematic review. *Croat Med J* 2019;60:309–15.
7. Emmett M, Mehta RH, Suzuki T. Predicting death in patients with acute type A aortic dissection. *Circulation* 2002;106:e224.
8. Tolenaar JL, Froehlich W, Jonker FH. Predicting in-hospital mortality in acute type B aortic dissection: evidences from IRAD. *Circulation* 2013;128.
9. Vrsalovic M, Zeljkovic I, Presecki AV, *et al*. C-Reactive protein, not cardiac troponin T, improves risk prediction in hypertensives with type A aortic dissection. *Blood Press* 2015;24:212–6.
10. Guo T, Zhou X, Zhu A, *et al*. The role of serum tenascin-C in predicting in-hospital death in acute aortic dissection. *Int Heart J* 2019;60:919–23.
11. Li G, Wu X-W, Lu W-H, *et al*. High-Sensitivity cardiac troponin T: a biomarker for the early risk stratification of type-A acute aortic dissection? *Arch Cardiovasc Dis* 2016;109:163–70.
12. Ohlmann P, Faure A, Morel O, *et al*. Diagnostic and prognostic value of circulating D-dimers in patients with acute aortic dissection. *Crit Care Med* 2006;34:1358–64.
13. Wen D, Du X, Dong J-Z, *et al*. Value of D-dimer and C reactive protein in predicting inhospital death in acute aortic dissection. *Heart* 2013;99:1192–7.
14. Wen D, Jia P, Du X, *et al*. Value of N-terminal pro-brain natriuretic peptide and aortic diameter in predicting in-hospital mortality in acute aortic dissection. *Cytokine* 2019;119:90–4.
15. Debray TPA, Damen JAAG, Snell KIE, *et al*. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;356:i6460.
16. Hemingway H, Croft P, Perel P, *et al*. Prognosis research strategy (progress) 1: a framework for researching clinical outcomes. *BMJ* 2013;346:e5595.
17. Geersing G-J, Bouwmeester W, Zuithoff P, *et al*. Search filters for finding prognostic and diagnostic prediction studies in medline to enhance systematic reviews. *PLoS One* 2012;7:e32844.
18. Pencina MJ, D'Agostino RB. Evaluating discrimination of risk prediction models: the C statistic. *JAMA* 2015;314:1063–4.
19. Moons KGM, de Groot JAH, Bouwmeester W, *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the charms checklist. *PLoS Med* 2014;11:e1001744.
20. Wolff RF, Moons KGM, Riley RD, *et al*. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:51–8.
21. Tan J, Qi Y, Liu C, *et al*. The use of rigorous methods was strongly warranted among prognostic prediction models for obstetric care. *J Clin Epidemiol* 2019;115:98–105.
22. Sahle BW, Owen AJ, Chin KL, *et al*. Risk prediction models for incident heart failure: a systematic review of methodology and model performance. *J Card Fail* 2017;23:680–7.
23. Janssen KJM, Donders ART, Harrell FE, *et al*. Missing covariate data in medical research: to impute is better than to ignore. *J Clin Epidemiol* 2010;63:721–7.
24. Meinshausen N, Bühlmann P. Stability selection. *J R Statist Soc B* 2010;72:417–73.
25. Vickers AJ, Cronin AM. Traditional statistical methods for evaluating prediction models are uninformative as to clinical value: towards a decision analytic framework. *Semin Oncol* 2010;37:31–8.
26. Pavlou M, Ambler G, Seaman SR, *et al*. How to develop a more accurate risk prediction model when there are few events. *BMJ* 2015;351:h3868.
27. Liu J, Sun L-L, Wang J, *et al*. The relationship between fibrinogen and in-hospital mortality in patients with type A acute aortic dissection. *Am J Emerg Med* 2018;36:741–4.
28. Zindovic I, Luts C, Bjursten H, *et al*. Perioperative hyperlactemia is a poor predictor of outcome in patients undergoing surgery for acute type-A aortic dissection. *J Cardiothorac Vasc Anesth* 2018;32:2479–84.
29. Oz K, Iyigun T, Karaman Z, *et al*. Prognostic value of neutrophil to lymphocyte ratio and risk factors for mortality in patients with Stanford type A aortic dissection. *Heart Surg Forum* 2017;20:119–23.
30. Feng W-Z, Zhou J-Q, Yu G-M, *et al*. Association of serum cystatin C levels with mortality in patients with acute type A aortic dissection. *Oncotarget* 2017;8:101103–11.
31. Karakoyun S, Gürsoy MO, Akgün T, *et al*. Neutrophil-Lymphocyte ratio may predict in-hospital mortality in patients with acute type A aortic dissection. *Herz* 2015;40:716–21.
32. Liu J, Sun L-L, Wang J, *et al*. Blood urea nitrogen in the prediction of in-hospital mortality of patients with acute aortic dissection. *Cardiol J* 2018;25:371–6.
33. Bennett JM, Wise ES, Hocking KM, *et al*. Hyperlactemia predicts surgical mortality in patients presenting with acute Stanford type-A aortic dissection. *J Cardiothorac Vasc Anesth* 2017;31:54–60.
34. Lafçi G, Ciçek Ömer Faruk, Uzun HA, *et al*. Relationship of admission neutrophil-to-lymphocyte ratio with in-hospital mortality in patients with acute type I aortic dissection. *Turk J Med Sci* 2014;44:186–92.
35. Zhang R, Chen S, Zhang H, *et al*. Biomarkers investigation for in-hospital death in patients with Stanford type A acute aortic dissection. *Int Heart J* 2016;57:622–6.
36. Li G, Zhao L, Ma Y. Platelet count to lymphocyte count ratio may predict mortality in Stanford type B acute aortic dissection. *International Journal of Clinical and Experimental Medicine* 2019;12:1922–8.
37. Zhang Y, Xu X, Lu Y, *et al*. Preoperative uric acid predicts in-hospital death in patients with acute type A aortic dissection. *J Cardiothorac Surg* 2020;15:21.
38. Bedel C, Selvi F. Association of platelet to lymphocyte and neutrophil to lymphocyte ratios with in-hospital mortality in patients with type A acute aortic dissection. *Braz J Cardiovasc Surg* 2020;34:694–8.
39. Gong M, Wu Z, Guan X, *et al*. Comparison of prognostic ability of perioperative myocardial biomarkers in acute type A aortic dissection. *Medicine* 2019;98:e17023.
40. Zhang J, Jiang Y, Gao C, *et al*. Risk factors for hospital death in patients with acute aortic dissection. *Heart, Lung and Circulation* 2015;24:348–53.
41. Ghoreishi M, Wise ES, Croal-Abrahams L, *et al*. A novel risk score predicts operative mortality after acute type A aortic dissection repair. *Ann Thorac Surg* 2018;106:1759–66.
42. Centofanti P, Flocco R, Ceresa F, *et al*. Is surgery always mandatory for type A aortic dissection? *Ann Thorac Surg* 2006;82:1658–64.
43. Santini F, Montalbano G, Casali G, *et al*. Clinical presentation is the main predictor of in-hospital death for patients with acute type A aortic dissection admitted for surgical treatment: a 25 years experience. *Int J Cardiol* 2007;115:305–11.
44. Rampoldi V, Trimarchi S, Eagle KA, *et al*. Simple risk models to predict surgical mortality in acute type A aortic dissection: the International registry of acute aortic dissection score. *Ann Thorac Surg* 2007;83:55–61.
45. Leontyev S, Légaré J-F, Borger MA, *et al*. Creation of a Scorecard to predict in-hospital death in patients undergoing operations for acute type A aortic dissection. *Ann Thorac Surg* 2016;101:1700–6.
46. Zhang J, Cheng B, Yang M, *et al*. Predicting in-hospital death in patients with type B acute aortic dissection. *Medicine* 2019;98:e16462.
47. Macrina F, Puddu PE, Sciangula A, *et al*. Long-Term mortality prediction after operations for type A ascending aortic dissection. *J Cardiothorac Surg* 2010;5:42.
48. Macrina F, Puddu PE, Sciangula A, *et al*. Artificial neural networks versus multiple logistic regression to predict 30-day mortality after operations for type A ascending aortic dissection. *Open Cardiovasc Med J* 2009;3:81–95.
49. Ge Y, Sun L, Zhu J, *et al*. Can EuroSCORE II predict the mortality and length of intensive care unit stay after total aortic arch replacement with stented elephant trunk implantation for DeBakey type I aortic dissection? *Thorac Cardiovasc Surg* 2013;61:564–8.
50. Yu P-J, Cassiere HA, Kohn N, *et al*. Utility of established risk models to predict surgical mortality in acute type-A aortic dissection. *J Cardiothorac Vasc Anesth* 2016;30:39–43.