Review Article

Predicting time to asystole following withdrawal of life-sustaining treatment: a systematic review

C. Nicolson,[1,2] A. Burke,[3] D. Gardiner,[4,5] D. Harvey,[4,5] L. Munshi,[6,7] M. Shaw,[8,9] A. Tsanas,[10] N. Lone[10,11] and K. Puxty[12,13]

1 Honorary Clinical Fellow, Department of Critical Care, NHS Greater Glasgow and Clyde, Glasgow Royal Infirmary, Glasgow, UK
2 PhD Researcher, School of Informatics, University of Edinburgh, Edinburgh, UK
3 Medical Student, Edinburgh Medical School, University of Edinburgh, Edinburgh, UK
4 Consultant, Critical Care, Nottingham University Hospitals NHS Trust, Derby Road, Nottingham, UK
5 National Clinical Lead, NHS Blood and Transplant, Watford, UK
6 Assistant Professor, Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, ON, Canada
7 Consultant, Department of Medicine, Sinai Health System and University Health Network, Toronto, ON, Canada
8 Principal Clinical Physicist, Department of Clinical Physics & Bioengineering, NHS Greater Glasgow and Clyde, Glasgow Royal Infirmary, Glasgow, UK
9 Honorary Senior Lecturer, School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK
10 Professor, Usher Institute, University of Edinburgh, Edinburgh, UK
11 Consultant, Department of Critical Care, NHS Lothian, Royal Infirmary of Edinburgh, Edinburgh, UK
12 Consultant, Department of Critical Care, NHS Greater Glasgow and Clyde, Glasgow Royal Infirmary, Glasgow, UK
13 Associate Professor, School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK

Summary

The planned withdrawal of life-sustaining treatment is a common practice in the intensive care unit for patients where ongoing organ support is recognised to be futile. Predicting the time to asystole following withdrawal of life-sustaining treatment is crucial for setting expectations, resource utilisation and identifying patients suitable for organ donation after circulatory death. This systematic review evaluates the literature for variables associated with, and predictive models for, time to asystole in patients managed on intensive care units. We conducted a comprehensive structured search of the MEDLINE and Embase databases. Studies evaluating patients managed on adult intensive care units undergoing withdrawal of life-sustaining treatment with recorded time to asystole were included. Data extraction and PROBAST quality assessment were performed and a narrative summary of the literature was provided. Twenty-three studies (7387 patients) met the inclusion criteria. Variables associated with imminent asystole (<60 min) included: deteriorating oxygenation; absence of corneal reflexes; absence of a cough reflex; blood pressure; use of vasopressors; and use of comfort medications. We identified a total of 20 unique predictive models using a wide range of variables and techniques. Many of these models also underwent secondary validation in further studies or were adapted to develop new models. This review identifies variables associated with time to asystole following withdrawal of life-sustaining treatment and summarises existing predictive models. Although several predictive models have been developed, their generalisability and performance varied. Further research and validation are needed to improve the accuracy and widespread adoption of predictive models for patients managed in intensive care units who may be eligible to donate organs following their diagnosis of death by circulatory criteria.

....

Correspondence to: C. Nicolson
Email: craig.nicolson@ed.ac.uk

Introduction

A common mode of death in the intensive care unit (ICU) involves the planned withdrawal of life-sustaining treatment (WLST) after recognising the futility of ongoing organ support. This may involve terminating invasive ventilation or vasopressors while end-of-life comfort care is administered [1]. While there is some international variation, death can be confirmed after a minimum of 5 min observation following the onset of mechanical asystole (hereafter asystole) [2, 3]. The prediction of time to asystole following WLST is important for setting expectations for families, ICU resource utilisation and to guide the identification of patients suitable for organ donation after circulatory death (DCD). The heterogeneity in underlying conditions, in combination with variations in levels of organ support, makes it challenging to accurately predict this time. Current practice to identify imminent death relies on clinical judgement; however, the abilities of physicians to make reliable predictions in this area are limited [4, 5].

The DCD donation process is often complex, resource intensive and can be emotionally distressing for families [6], particularly if the donation is unable to proceed [7]. Successful donation is often prevented due to logistical challenges or the occurrence of prolonged time to asystole where the organs are damaged due to excessive warm ischaemic time. In the UK, 45% of unsuccessful DCD donations are attributed to this prolonged time period [8].

A variety of predictive tools and models have been developed [4, 5, 9–19] which can provide support for decision-making in this area. Whilst some prediction models initially appear to perform well, they are often not appropriately validated and when external validation has been undertaken, their performance typically does not generalise well [20–22]. A lack of standardisation of the variables recorded and the specifics of the withdrawal process makes the transfer and shared use of developed tools challenging. These problems underline the unmet clinical need for the development of clinical decision support tools capitalising on data, which in this context could support widespread adoption and deployment of time to asystole prediction models in DCD. The aim of this systematic review was to evaluate the literature for variables associated with, and predictive models for, time to asystole in patients managed on ICUs and who are eligible for DCD.

Methods

Following registration with PROSPERO [23] we searched MEDLINE (inception to 11 May 2022) and Embase (inception to 11 May 2022). The searches combined Medical Subject Headings (MeSH), appropriate controlled vocabulary and keywords for time, death and withdrawal (online Supporting Information Appendices S1 and S2). We explored the reference lists of all included studies and prior review studies for further inclusions. Clinical experts were consulted to evaluate the list of included studies for omissions identified through their knowledge of the field. Conference abstracts, poster abstracts, letter responses and letters to editors were excluded. To meet inclusion criteria, we required studies to evaluate an adult population in an ICU environment who underwent WLST and had an associated time to asystole recorded. Life-sustaining treatment was defined as ventilation (invasive or non-invasive) or haemodynamic support. Measurement of time from WLST to death or asystole was also necessary for inclusion. Studies that did not evaluate potential predictive factors or models in relation to this measurement were excluded. Only English language studies were included.

Two reviewers (CN and AB) independently reviewed all titles and abstracts identified from the literature searches. Potentially eligible studies underwent duplicate full-text review. During both processes, we resolved disagreements through a third reviewer (KP).

We extracted data from studies using customised spreadsheets, with key population characteristics such as age, ICU type and methods of withdrawal recorded. The focus of the studies was summarised as a focus on one of: variable evaluation; predictive model development; or a mixture of both. Other study design aspects such as the methods of withdrawal and outcomes assessed were also recorded. The performance metrics of any evaluated predictive factors or models were recorded.

We undertook quality assessment of predictive model development and validation using the PROBAST (Prediction model Risk Of Bias ASsessment Tool) [24] which was designed to assess the risk of bias and applicability of

diagnostic and prognostic prediction model studies. Given the lack of standardisation of the withdrawal process, heterogeneous populations and variables measured at time of WLST, and variation in the outcome measure, we did not undertake data pooling or meta-analysis. Therefore, the analysis consisted of tabulation of study characteristics and performance metrics with narrative summarisation of the literature.

Results

The initial search returned a total of 2418 studies to be further screened following the removal of duplicates (Fig. 1). This produced 71 studies for full-text review with an additional paper from reference screening and from expert input. Full-text review identified 23 studies (7387 patients) for inclusion in data extraction and analysis.

The DCD status of the populations was relatively evenly split with 11 DCD eligible patient populations [9, 11, 12, 14, 18–22, 25, 26], 11 general populations not restricted by DCD eligibility [4, 5, 10, 13, 15, 17, 27–31] and one study evaluating a DCD eligible population [16] that was a subset of a previously evaluated general population [15]. Median

ages of the patient populations ranged from 41 to 66 years. Mixed ICU patient populations were the focus of 15 studies [4, 5, 9, 10, 12, 14, 15, 19, 20, 25–27, 29–31] with a total of 5131 patients. Neurointensive care patient populations were the focus of four studies [17, 21, 22, 28] with a total of 1181 patients. The remaining three studies [11, 13, 18] did not specify the type of ICU environment from which the patients originated, and totalled 757 patients. These study characteristics are tabulated in the online Supporting Information Table S1.

Practices for WLST varied across the studies. Mechanical ventilation of the patient's lungs was stopped at the point of WLST in all studies, with the cessation of vasoactive drugs in the majority. The specific process of withdrawal, including details surrounding the administration of comfort care medicines was not typically detailed, with only five studies [11, 14, 19, 20, 25] specifying that withdrawal of all active treatments was simultaneous.

Most studies evaluated the likelihood of asystole within 60 min as an outcome, with the exception of five studies [18, 26, 27, 29, 31]. The proportion of patients who became asystolic within 60 min ranged from 44–76%. Eight studies
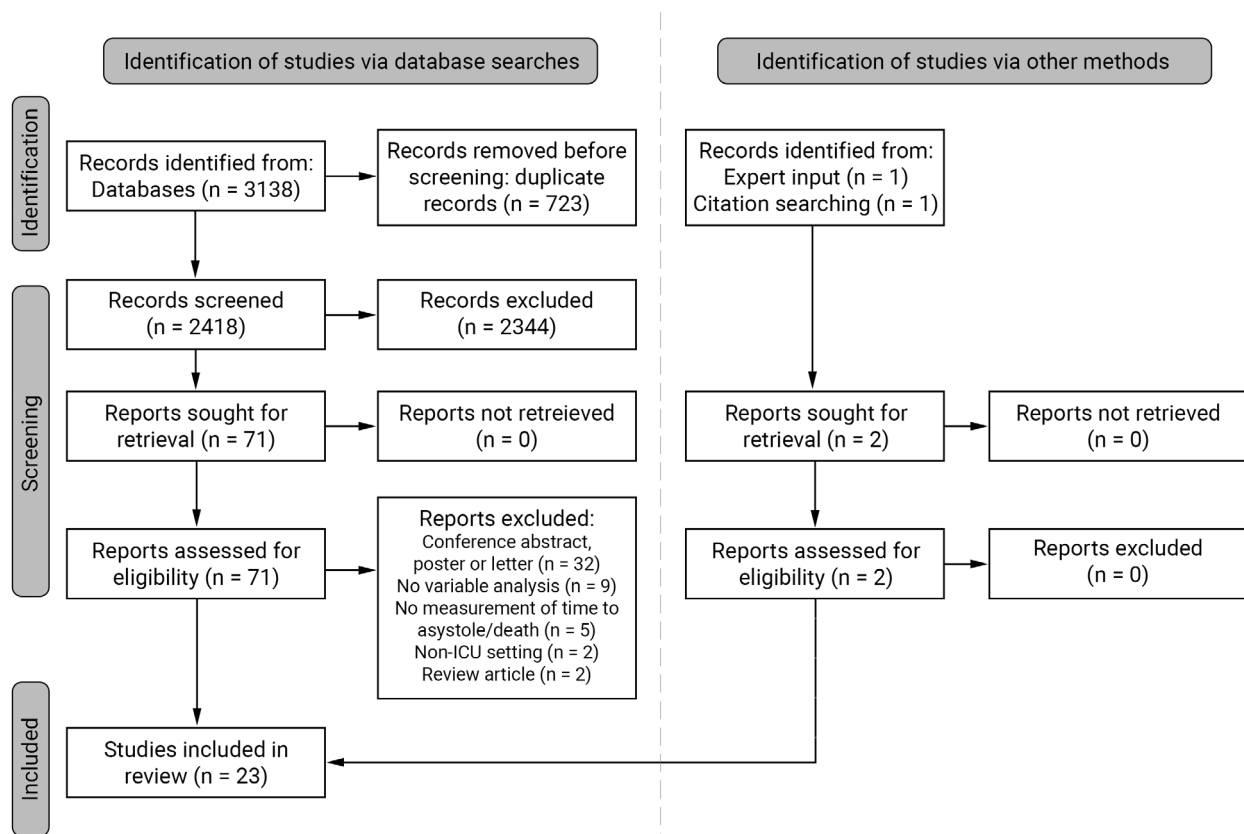

Figure 1 Study flow diagram of study identification and inclusion.

3

[4, 9, 13, 14, 17, 19, 21, 26] evaluated the likelihood of asystole within 120 min, which occurred in 54–91% of patients. Full outcomes assessed are shown in Table S1.

Seven studies [13, 25–27, 29–31] focused on the evaluation of variables associated with their primary outcome and did not derive or validate any predictive models. These studies typically used p value cut-off regression techniques. The variables found to be statistically associated with asystole within 60 min using multivariable analysis are detailed in Table 1 including odds and/or hazard ratios with confidence intervals.

Deteriorating oxygenation was associated with imminent asystole (<60 min) in six analyses represented by SaO_2:F_IO_2 ratio < 230 and PaO_2 < 9.6 kPa [10]; oxygenation index [13, 20, 28] (which assesses the degree of ventilatory support required to maintain oxygenation [32]); F_IO_2 > 70% [30]; and increased F_IO_2 (>50%) [25].

The absence of corneal reflexes and a cough reflex were found to be associated with imminent asystole in four analyses [13, 17, 20, 28]. Additionally, an absent or extensor motor response was associated with three of these analyses [13, 17, 28].

Blood pressure measurements and the use of vasopressors were associated with time to asystole. However, their methods of evaluation varied between the studies: lower diastolic blood pressure [10]; lower systolic blood pressure [15]; lower mean arterial pressure [31]; higher vasopressor dose (adrenaline, noradrenaline or phenylephrine > 0.2 $\mu g.kg^{-1}.min^{-1}$) [10]; vasopressor use prior to withdrawal [27]; and vasopressor use within 12 h of withdrawal [30] were associated with time to asystole.

Two analyses identified the use of comfort medications following WLST as being inversely associated, where the use of comfort medications reduced the odds of asystole within 60 min. In the study by DeVita et al. [10] the use of comfort medications during the first hour after WLST reduced the odds of imminent death during that time. In the study by Kotsopoulos et al. [20], the use of midazolam and dose of morphine administered after WLST were also associated with reduced odds of asystole within 60 min. We agree with Devita et al. [10] that this paradoxical finding warrants a prospective trial to evaluate causality. The final associated variable identified in multiple analyses was positive end-expiratory pressure [15, 31].

The derivation or modification of predictive models was reported in eight studies [4, 11, 13–17], external validation of an existing model was undertaken in three [20–22] and both derivation and external validation were undertaken in five [5, 9, 10, 12, 19] (Table 2). Most of the predictive model studies (15/16) included the evaluation of

models for prediction of asystole within 60 min whilst some included evaluation of asystole within 120 min (9/16) or other time ranges (2/16). In total 15 original models were reported with a further five models derived by augmenting or adjusting these.

Two of the original models evaluated were developed using clinical experience and expert consensus without the reported use of statistical techniques [9, 10]. Of the remaining original models, two used a classification and regression tree (CART) [10, 16], two used Cox regression analysis [17, 25] and seven used other forms of multivariable regression analysis [5, 11–15, 19]. The final two original models used random survival forests [4] and a light gradient boosting machine [19] respectively.

At the point of derivation or modification validation procedures were varied, with five studies evaluating model performance against the same cohort used for model fitting [5, 9, 10, 14, 19], with only one of these using cross-validation to attempt to mitigate the impact of overfitting [19]. Four models were validated at the point of derivation by randomly splitting the cohort into a training and a testing set [4, 12, 15, 16], with a further four validated using an external cohort. One model was validated using a prospective cohort and one was validated using both external and prospective cohorts.

Sixteen instances of secondary validation (validation of a distinct model by another group) were observed across six studies [5, 12, 19–22] with three of these studies solely attempting to validate previously derived models without any new model derivation or modification of existing models.

The University of Wisconsin DCD tool (UWDCD) [9] is a scoring system developed using clinical experience to identify relevant clinical and demographic patient characteristics. As part of the tool, spontaneous respiratory efforts, tidal volume, negative inspiratory force and oxygen saturation are measured during a 10-min period of ventilator disconnection. The tool thresholds these measurements alongside the number of drugs used for blood pressure support, patient age and airway type. Initial validation using 43 patients showed a sensitivity of 0.87 and specificity of 0.80 for asystole within 60 min of WLST [9]. During this validation, the authors also explored the inclusion of BMI, yielding a sensitivity of 0.84 and specificity of 0.85. Two studies externally validated the tool that included BMI in larger populations (81 and 219 patients) where the performance of predicting asystole within 60 min did not generalise sufficiently well, reporting sensitivities of 0.42 and 0.45 and specificities of 0.61 and 0.49 [5, 21].

The United Network for Organ Sharing tool (UNOS) [10] was developed using committee consensus and was

Table 1 Variables associated with asystole within 60 min of withdrawal of life-sustaining treatment using multivariable analysis. Alternative tabulation of risk ratios for the ten most frequently identified predictors available in online Supporting Information Table S3. Values are OR (95%CI) or hazard ratio (95%CI).

Publication	Risk factor	Odds ratio	p value
DeVita et al. [10]	GCS 3	2.83 (1.79–4.46)	<0.001
	$SaO_2/F_IO_2 < 230$	1.78 (1.09–2.90)	<0.05
	Peak inspiratory pressure > 35 cm H_2O	2.58 (1.49–4.48)	<0.001
	Respiratory rate off ventilator < 8 breaths.min^{-1}	6.01 (2.29 –4.48)	<0.001
	Diastolic blood pressure (per 10 mmHg)	0.80 (0.69–0.93)	<0.01
	$PaO_2 < 9.56$ kPa	3.10 (1.53–6.30)	<0.01
	Adrenaline, noradrenaline or phenylephrine > 0.2 $\mu g^{-1}.kg^{-1}.min^{-1}$	3.02 (1.75–5.21)	<0.001
	All treatments withdrawn within 10 min	8.55 (4.23–17.30)	<0.001
	Tracheal tube removed	2.28 (1.33–3.90)	<0.01
	Comfort medications given during first hour after WLST	0.35 (0.21–0.59)	<0.001
Yee et al. [28]	Absent corneal reflex	4.24 (1.57–11.5)	0.0045
	Absent cough reflex	4.47 (1.93–10.3)	0.0005
	Extensor or absent motor response	2.83 (1.01–7.91)	0.048
	Oxygenation index > 4.2	3.36 (1.33–8.50)	0.0105
Rabinstein et al. [13]	Absent corneal reflex	2.67 (1.19–6.01)	0.02
	Absent cough reflex	4.16 (1.79–9.70)	<0.001
	Extensor or absent motor response	2.99 (1.22–7.34)	0.02
	Oxygenation index > 3.0	2.31 (1.10–4.85)	0.03
Wind et al. [14]	Controlled mechanical ventilation	2.50 (1.18–5.26)	0.02
Brieva et al. [15]	Physician opinion	8.44 (4.3–16.6)	<0.001
	pH (per 0.1 units)	0.67 (0.47–0.94)	0.03
	Systolic blood pressure (per mmHg)	0.99 (0.98–1.00)	0.01
	GCS (per unit of GCS)	0.85 (0.74–0.98)	0.03
	Positive end-expiratory pressure (per cmH_2O)	1.17 (1.06–1.30)	<0.01
	Analgesia	0.32 (0.15–0.70)	<0.01
Huynh et al. [30]	Vasopresssors within 12 h of MV withdrawal	2.05 (1.37–3.07)	-
	$FIO_2 > 70\%$	1.92 (1.24–2.98)	-
Kotsopoulos et al. [20]	Absent cough reflex	8.04 (5.51–11.7)	<0.001
	Absent corneal reflex	3.76 (2.65–5.32)	<0.001
	Dosage of morphine at WLST	0.88 (0.82–0.95)	0.002
	Midazolam use after WLST	0.15 (0.09–0.23)	<0.001
	Oxygenation index (per unit)	1.18 (1.17–2.80)	0.008

Publication	Risk factor	Hazard ratio	p value
Suntharalingam et al. [25]	Age 31–40 y	0.70 (0.38–1.28)	0.001
	Age 41–50 y	0.46 (0.29–0.76)	0.001
	Age > 50 y	0.37 (0.23–0.59)	0.001
	F_IO_2 (per %)	1.01 (1.00–1.02)	0.008
	Ventilation without pressure support	1.67 (1.16–2.41)	0.006
Cooke et al. [27]	Age (per 10 y)	0.95 (0.90–0.99)	-
	Female	0.86 (0.77–0.97)	-
	Non-white ethnicity	1.17 (1.01–1.35)	-
	Number of organ failures	1.11 (1.04–1.19)	-
	Surgical service	1.29 (1.06–1.56)	-

(continued)

Table 1 (continued)

Publication	Risk factor	Hazard ratio	p value
	Vasopressors	1.67 (1.49–1.88)	-
	Intravenous fluids	1.16 (1.01–1.32)	-
He et al. [17]	Pupils (normal vs. dilated)	0.34 (0.20–0.57)	<0.001
	Pupillary reflex (fixed vs. brisk)	0.53 (0.29–0.95)	0.034
	Corneal reflex (present vs. absent)	0.32 (0.16–0.64)	0.001
	Cough reflex (present vs. absent)	0.45 (0.27–0.74)	0.002
	Motor response (normal vs. absent)	0.41 (0.28–0.60)	<0.001
	Cisterna ambiens (normal vs. absent)	0.19 (0.08–0.43)	<0.001
	Swirl sign (present vs. absent)	0.36 (0.24–0.54)	<0.001
	Brain herniation (present vs. absent)	0.44 (0.31–0.64)	<0.001
	Intraventricular haemorrhage (present vs. absent)	0.28 (0.19–0.42)	<0.001
Long et al. [31]	Diabetes (present vs. absent)	1.75 (1.25–2.45)	0.001
	PEEP (cmH$_2$O)	1.07 (1.04–1.11)	<0.001
	Static pressure (cmH$_2$O)	1.03 (1.01–1.04)	0.006
	Tracheal extubation	1.41 (1.06–1.86)	0.017
	Non-invasive MAP (mmHg)	0.98 (0.97–0.99)	0.004

	Predictive variable	No. of studies identifying variable
Most frequently identified predictors	Oxygenation	6
	Corneal reflex	4
	Cough reflex	4
	Motor response	3
	Blood pressure	3
	Use of vasopressors	3
	Use of comfort medications	2
	PEEP	2
	GCS	2

GCS, Glasgow Coma Scale; MV, mechanical ventilation; WLST, withdrawal of life-sustaining treatment; PEEP, positive end-expiratory pressure; MAP, mean arterial pressure.

evaluated in three studies. It is a criteria-based scoring tool with higher scores corresponding to a higher probability of asystole within 60 min. As with the UWDCD tool, the original criteria use four measurements during a period of ventilator disconnection in addition to seven further clinical parameters. Two evaluations of the full criteria aiming to assess prediction of asystole within 60 min produced an area under the receiver operating characteristic curve (AUROC) of 0.53 [21] and a positive predictive value (PPV) of 0.63 [10]. The evaluation by Coleman et al. used a modified UNOS tool where the period of ventilator disconnection was omitted and replaced with definitions of ventilator dependence and oxygen disruption; this was found to give a sensitivity of 0.61 and specificity of 0.84 [5].

Devita et al. [10] developed a predictive tool using CART model analysis which took GCS $\geq$ 4, SaO$_2$/F$_1$O$_2$ $\geq$ 230 and positive inspiratory pressure $\geq$ 35 cm H$_2$O as inputs. The performance was evaluated directly on the derivation cohort, giving a sensitivity of 0.75 and specificity of 0.73 but no external validation has been undertaken.

The Hunter New England Area Composite score [5] was developed using logistic regression and included ventilatory dependence with a measure of oxygenation disruption. It achieved sensitivity of 0.56 and specificity of 0.13 on the derivation cohort. The inclusion of a systolic blood pressure threshold (<100 mmHg) led to sensitivity of 0.39 and specificity of 0.96. No external validation of this score has been undertaken.

Table 2 Developed and validated predictive models and performance metrics. The full list of variables used in each model and their corresponding ranges and/or thresholds are reported in online Supporting Information Table S4.

	Model number	Type of model	Time to asystole; min	AUROC or Harrell's c-index* (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)
Lewis et al. [9]	1	Validation**	<60	-	0.87	0.80	0.83	0.84
			<120	-	0.79	0.64	0.82	0.60
	2	Development	<60	-	0.84	0.85	0.91	0.73
			<120	-	0.79	0.56	0.85	0.45
DeVita et al. [10]	3	Validation**	<60	-	-	-	0.63	-
	4	Development		-	0.79	0.63	0.63	0.78
	5	Development		-	0.75	0.73	0.69	0.78
Coleman et al. [5]	1	Validation	<60	-	0.42	0.76	0.58	0.62
	2	Validation		-	0.61	0.84	0.76	0.73
	6	Development		-	0.56	0.13	0.34	0.27
	7	Development		-	0.39	0.96	0.88	0.66
Davila et al. [11]	8	Development	<60	0.83 (0.76–0.90)	0.86	0.80	-	-
de Groot et al. [12]	9	Validation***	<60	0.75 (0.63–0.87)	-	-	-	-
				0.73 (0.61–0.85)	-	-	-	-
	10	Development		0.74 (0.66–0.87)	-	-	-	-
Rabinstein et al. [13]	9	Development	<60	-	0.81	0.73	0.75	0.79
Wind et al. [14]	11	Development	<60	0.73 (0.66–0.82)	0.70	0.74	-	-
			<120	0.78 (0.69–0.86)	0.84	0.53	-	-
Brieva et al. [15]	12	Development	<60	0.78 (0.73–0.83)	0.82 (0.75–0.87)	0.59 (0.52–0.66)	0.68 (0.62–0.74)	0.75 (0.67–0.81)
	12#			0.84 (0.80–0.88)	0.84 (0.78–0.89)	0.72 (0.65–0.79)	0.77 (0.71–0.82)	0.81 (0.74–0.86)
Brieva et al. [16]	13	Development	<60	-	0.66 (0.55–0.76)	0.77 (0.66–0.86)	0.77 (0.65–0.86)	0.66 (0.55–0.76)
	13#			-	0.82 (0.73–0.90)	0.76 (0.64–0.85)	0.80 (0.70–0.87)	0.79 (0.68–0.88)
He et al. [17]	14	Development	<60	0.94* (0.91–0.97)	-	-	-	-
			<120	0.95* (0.92–0.98)	-	-	-	-
			<60	0.99* (0.97–1.00)	-	-	-	-
			<120	0.84* (0.72–0.96)	-	-	-	-
Kotsopoulos et al. [20]	10	Validation	<60	0.86 (0.77–0.95)	-	-	-	-
	8	Validation		0.80 (0.70–0.90)	-	-	-	-
	11	Validation		0.62 (0.49–0.76)	-	-	-	-
	15	Validation		0.63*	-	-	-	-
Xu et al. [21]	1	Validation	<60	0.45 (0.36–0.53)	-	-	-	-
			<120	0.49 (0.38–0.61)	-	-	-	-
	3	Validation	<60	0.53 (0.46–0.59)	-	-	-	-
			<120	0.51 (0.43–0.60)	-	-	-	-
	9	Validation	<60	0.69 (0.62–0.77)	-	-	-	-
			<120	0.73 (0.63–0.83)	-	-	-	-
	14	Validation	<60	0.88 (0.83–0.93)	-	-	-	-
			<120	0.86 (0.77–0.94)	-	-	-	-
Scales et al. [4]	16	Development	<60	0.79 (0.71–0.87)	0.87 (0.71–0.97)	0.7 (0.53–0.85)	0.75 (0.67–0.84)	0.83 (0.71–0.95)
			<120	0.8 (0.72–0.88)	0.78 (0.64–0.91)	0.77 (0.58–0.90)	0.86 (0.79–0.93)	0.66 (0.55–0.81)
Nijhoff et al. [22]	9	Validation	<60	0.77 (0.71–0.83)	0.88	0.46	0.36	0.92
			<120	0.80 (0.74–0.86)	-	-	-	-
	10	Validation	<60	0.75 (0.68–0.81)	0.83	0.51	0.38	0.89
			<120	0.83 (0.77–0.88)	-	-	-	-
Okahara et al. [18]	17	Development	<90	-	1	-	0.56	-
Kotsopoulos et al. [19]	10	Validation	<60 m	0.74 (0.70–0.79)	-	-	-	-
				0.86 (0.77–0.95)	-	-	-	-
	8	Validation		0.70 (0.65–0.75)	-	-	-	-
				0.80 (0.70–0.90)	-	-	-	-
	15	Validation		0.63 (0.58–0.68)	-	-	-	-

(continued)

Table 2 (*continued*)

	Model number	Type of model	Time to asystole; min	AUROC or Harrell's c-index* (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)
				0.63	-	-	-	-
	18	Development	<60 m	0.80 (0.69–0.91)	-	-	-	-
	19	Development	<120 m	0.82 (0.73–0.92)	-	-	-	-
	20	Development	<60 m	0.79	-	-	-	-
			<120 m	0.81	-	-	-	-

*, **First published validation of a tool without initial published derivation; ***Validation of the variable combination identified by Yee et al. [28] before evaluation; #Plus expert opinion.

Table 3 Tabular Prediction model Risk Of Bias ASessment Tool (PROBAST) results for risk of bias and applicability across domains.

Study	Risk of bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
Lewis et al. [9]	+	+	+	-	+	-	+	-	-
DeVita et al. [10]	+	+	?	-	-	-	+	-	-
Coleman et al. [5]	+	+	+	-	-	+	+	-	-
Davila et al. [11]	+	+	+	-	-	+	+	-	-
de Groot et al. [12]	?	+	?	-	-	+	+	-	-
Rabinstein et al. [13]	+	+	+	-	-	+	+	-	-
Wind et al. [14]	+	+	+	-	+	+	+	-	-
Brieva et al. [15]	+	+	+	-	-	+	+	-	-
Brieva et al. [16]	+	+	+	-	+	+	+	-	+
He et al. [17]	+	+	+	+	-	-	+	+	-
Kotsopoulos et al. [20]	+	+	+	+	+	+	+	+	+
Xu et al. [21]	-	+	+	+	-	-	+	?	-
Scales et al. [4]	?	+	+	+	-	+	+	+	-
Nijhoff et al. [22]	+	+	+	+	+	+	+	+	+
Okahara et al. [18]	+	+	+	-	-	-	-	-	-
Kotsopoulos et al. [19]	+	+	+	+	+	+	+	+	+

+ indicates low risk of bias/low concern regarding applicability; - indicates high risk of bias/high concern regarding applicability; and ? indicates unclear risk of bias/unclear concern regarding applicability.

Davila et al. [11] used multivariable regression analysis to develop a tool that took the dichotomised variables of vasoactive drug use, age ≤ 40 y and presence of gag reflex. The original authors validated the tool using a prospective cohort to give an AUROC of 0.83. External validation across three cohorts yielded an AUROC of 0.80, 0.70 and 0.80 [19, 20].

The circulatory death in patients in neurocritical state (DCD-N) tool is based on corneal reflex, cough reflex, motor response and oxygenation index [28]. The authors validated the model through fitting with a prospective cohort [13], finding a sensitivity of 0.81 and 0.73. Three studies externally validated this model giving an AUROC of 0.75 [12], 0.69 [21] and 0.77 [22]. In addition to externally validating this model, de Groot et al. [12] also proposed a modified model using a

continuous oxygen index. This modified model gave an AUROC of 0.74 using an internal validation cohort and was externally validated across four cohorts giving an AUROC of 0.75, 0.86, 0.74 and 0.86 [19, 20, 22].

Wind et al. [14] used logistic regression to develop a model that used the dichotomised presence of: controlled mode ventilation; noradrenaline use; cardiovascular co-morbidity; brainstem reflexes; and neurologic deficit. This achieved an AUROC of 0.73 within the derivation cohort which fell to 0.62 in a subsequent external validation cohort [20].

Brieva et al. developed models using logistic regression in a general population [15] and later, using CART model analysis in a DCD subset of this patient population [16]. All of these models used ranges of positive

end-expiratory pressure, spontaneous respiratory rate, GCS and systolic blood pressure, with the inclusion of ICU specialist opinion used to define a second model in both populations. Using internal validation cohorts, the models achieved sensitivities of 0.82 and 0.66 and specificities of 0.59 and 0.77 in the general and DCD patient populations, respectively. The inclusion of ICU specialist opinion in these models improved performance to sensitives of 0.84 and 0.82, and specificities of 0.72 and 0.76 respectively.

The C-DCD model is a nomogram that was developed using Cox regression analysis to identify 10 variables for inclusion (Table S4) [17]. The authors used two validation cohorts (external and prospective) to evaluate the model and report an AUROC of 0.94 and 0.99. A subsequent external validation demonstrated slightly worse performance with an AUROC of 0.88 [21].

The first model to incorporate time series rather than instantaneous data into a prediction model was developed by Scales et al. [4]. Here random survival forests used physician predictions alongside a series of variability features of blood pressure and heart rate time series. An internal validation cohort demonstrated an AUROC of 0.79 using this model. There has been no external validation of this model.

Finally, alongside the validation of two previous models, Kotsopoulos et al. evaluated two new models, one using the least absolute shrinkage and selection operator (LASSO) to select variables in a standard linear least squares setting, and another using a light gradient boosting machine (LightGBM) which accounts for more convoluted nonlinear statistical relationship between the variables and the outcome [19]. The regression model which was built using LASSO achieved an AUROC of 0.80 in their external validation cohort and the model that was built using LightGBM achieved an AUC of 0.79 in 10-fold cross-validation.

We assessed the quality of the studies reporting model development or validation using PROBAST. The full results are available in online Supporting Information Table S2 with tabular results summarised in Table 3.

When considering all domains using the PROBAST assessment criteria, 10 studies were judged to be at high risk of bias, five were considered at low risk of bias and one was at an unclear risk of bias. Participant selection (domain 1) was considered to be low risk of bias in 13 studies, unclear in two studies due to inclusion/exclusion criteria, and high risk in one study due to participant selection. Predictors (domain 2) were judged to be low risk in all 16 studies, with consistent predictive variable definition and assessment. Outcome determination (domain 3) was found to be at low risk of bias in 14 studies and unclear in the remaining two due to lack of clarity of the withdrawal process. Analysis (domain 4) was found to be the most common source of risk of bias with a high result for 10 studies and a low result for the remaining six studies.

Data imbalance, and in particular lacking a sufficiently large number of participants who had died within the specified time span (60 min or 120 min), was identified as an issue in 13 of 16 studies. Potentially inappropriate dichotomisation of continuous variables was identified in 3 of 16 studies. The frequency of missing data was not reported in 10 studies and participants with missing data were excluded in two studies. Multiple imputation was used to mitigate missing data in three studies where missing data were < 5% and the methods used to manage missing data were not reported in the final study.

Variable selection based on univariable analysis was identified in eight studies, leading to the possible exclusion of variables that may have been important after adjustment for other variables or failing to account for information overlap (i.e. redundancy). Univariable analysis to determine a subset of predictive variables is known to be flawed and there is considerable research work in the machine learning community investigating advanced principled methods towards determining parsimonious variable subsets.

Appropriate model performance measures were used in 10 studies with this becoming more common in more recent publications. As defined within PROBAST, the consideration of overfitting and optimism were also more frequent in more recent publications, although only four model development studies accounted for these effects appropriately. Four studies published between 2003 and 2012 undertook model evaluation using the full derivation cohort which inevitably leads to overfitting and reported performances that would be unlikely to be verified in an external dataset.

When considering all domains, 12 studies were high concern for applicability and four studies were low concern for applicability (Table 3). Participant selection (domain 1) was found to be high concern for applicability in 10 of 16 studies, with this typically being due to either a non-DCD eligible population or restriction to specific or limited diagnoses. Predictors (domain 2) were considered to be of low concern for applicability in 11 studies and high concern in the remaining five. High concern for applicability was suggested by the use of a period of ventilator disconnection to generate predictive variables in two studies and the requirement for brain imaging in three studies.

Outcome applicability (domain 3) was generally good, with all but one study reporting the generally accepted 60 min threshold for asystole with many also reporting additional thresholds.

Discussion

This systematic review shows the ongoing challenges of developing reliable predictive tools for time to asystole following WLST. The heterogeneity of patient populations and variations in clinical practice continue to represent core challenges that impede validation and potential deployment.

We found that studies focussing on identifying variables associated with time to asystole have further strengthened the body of evidence behind the key predictors, with neurological and respiratory-related variables found repeatedly to be associated with time to asystole. In particular, the absence of a cough reflex or corneal reflexes were associated with a shorter time to asystole. It is possible that these are surrogate markers for the loss of other brainstem reflexes and may represent a population of patients who are apnoeic and will progress to cardiorespiratory arrest shortly after cessation of mandatory ventilation. We identified several tools that have been developed to predict time to asystole in <60 min although, in general, these did not perform well on external validation. The most promising tool to date was the C-DCD tool, which had an AUROC of 0.88 on external validation in a neurosurgical ICU setting. It has yet to undergo evaluation in a general ICU setting beyond the original paper.

We were able to systematically evaluate the quality of included studies using the PROBAST tool, with a focus on their applicability to the review question and their risk of bias. We found that over time studies have substantially improved in both these areas. Robust examples of primary and external validation have become more prevalent within the literature, which helps move developed predictive tools closer to clinical use. Despite this, there are still no reports within the literature of use of existing or developed predictive tools clinically. The reasons for this are not explicitly discussed or elaborated on.

This systematic review builds on a prior review by Munshi et al. [33] who evaluated the literature in this area up to 2014 with similar inclusion criteria. Since publication of this review, several new time-to-asystole prediction tools have been developed. There has also been a greater focus on external validation of existing prediction tools, often evaluating several at the same time. The focus on time to asystole within 60 min as the primary outcome has persisted, although many studies have expanded this scope to include evaluations of models targeting time to asystole within 30 min, 120 min or 240 min. The ongoing exploration of predicting asystole within periods >60 min is important, given the evolving landscape of DCD [34]. The modelling methods used for predictions have become more diverse with the first use of random survival forests, variable selection using LASSO methods and gradient boosting machines explored. Additionally, one study incorporated time series data through the use of measures of vital sign variability as predictive variables [4] in contrast to the use of a single instantaneous set of variables as seen in all other models.

The authors of the previous systematic review concluded that differences in practice could have an important impact on time to asystole as many studies did not specify the simultaneous withdrawal of ICU treatments (e.g. inotropic support and tracheal extubation) at time of WLST [34]. This weakness remains in many of the more recent publications and likely reflects the variation in clinical practice in ICUs and the problems in modifying or standardising clinical protocols for observational studies.

The use of more rigorous prediction tool development processes, particularly with regards to variable selection, validation measures and model calibration, has successfully reduced the bias observed in recent studies. This helps to provide support to the results as well as the performance measures they report in the external validation of existing models.

The increased focus on this external validation of existing models allows us to compare their results with the reported model performance at derivation. Three models [11, 12, 25] were externally validated in more than one external cohort, with predictive models typically performing slightly worse on external validation. This may suggest that the heterogeneity of patient populations is limiting generalisability or that the initial model development process was susceptible to optimism or overfitting. However, the broad reproducibility of model performance through validation using several large populations is promising and provides evidence that supports the underlying models.

The suitability of the described models for use in clinical practice is dependent on their specific use, patients and environments. We observed minimal discussion of the importance of model performance metrics in real-world clinical situations. The trade-offs of acceptable specificity and sensitivity will be strongly impacted by the clinical environment, organ donation resources and number of potential donors evaluated. For example, in a context where there are many more potential donor candidates than the organ donation resources could manage, it may be important that the predictive model has a high specificity to maximise the effectiveness of the organ donation resources such as organ retrieval teams.

Several models showed promising performance on external validation; however, none could reliably predict imminent asystole with a level of accuracy that precludes error. Consideration for use in DCD protocols will depend on their current performance at predicting imminent asystole. The case for this could be further strengthened by organisational level validation of the model's reliability and applicability before deployment.

There are several limitations to this review which should be noted. First, the inclusion of mixed intensive care and neurointensive care populations, and the inclusion of DCD eligible and non-DCD eligible populations lead to wide heterogeneity in the study participants. This limits the potential for both meta-analysis and conclusions about specific populations as there may be wide differences between these populations. Second, the models remain heavily focused on an outcome of time to asystole within 60 min, which limits the applicability in situations where longer times to asystole would be clinically acceptable, such as DCD programmes focused on specific organs.

The ongoing progress in predictive variable identification and predictive modelling suggests that we may be getting closer to developing and validating models that could be deployed more broadly in clinical practice. Current exploration and potential use of models appears to be fractured and often on a local scale or limited to a specific organ retrieval programme. The development of more generalisable tools will rely on further large multi-centre or international studies using robust analysis procedures. It would be prudent for future efforts to develop predictive tools to consider all predictive variables identified within this review to aid in planning prospective data collection and variable selection in their analyses. The use of time series data, as partially demonstrated with measures of variability [4], may offer several advantages over typical instantaneous variables where the negative impact of spurious or outlying measurements has the potential to be minimised. Further exploration into the use of vital sign variability and full time series data for predictive modelling is warranted.

In conclusion, reliable prediction of time to asystole following WLST is an important area for clinical practice for DCD and vital for its expansion within and in addition to, existing programs. Accurate prediction would ensure donation is offered as an end-of-life decision where it is feasible, for the benefit of potential donors, their families, organ donation systems and recipients. This review provides an overview of the current progress in this area, highlighting key challenges and identifying areas of potential interest for further exploration. We emphasise the need to move towards standardisation of WLST practices through consensus, and we single out the potential of extracting clinically useful information from time series data (e.g. mining blood pressure and heart rate series), which has been hitherto largely neglected.

Acknowledgements

References

1. Avidan A, Sprung CL, Schefold JC, et al. Variations in end-of-life practices in intensive care units worldwide (Ethicus-2): a prospective observational study. *Lancet Respiratory Medicine* 2021; **9**: 1101–10.
2. Dhanani S, Hornby L, van Beinum A, et al. Resumption of cardiac activity after withdrawal of life-sustaining measures. *New England Journal of Medicine* 2021; **384**: 345–52.
3. Lomero M, Gardiner D, Coll E, et al. Donation after circulatory death today: an updated overview of the European landscape. *Transplant International* 2020; **33**: 76–88.
4. Scales NB, Herry C, Van Beinum A, et al. Predicting time to death after the withdrawal of life supporting therapies using variability analysis of vital signs waveform data. *Canadian Journal of Anesthesia* 2019; **66**: S96–S100.
5. Coleman NL, Brieva JL, Crowfoot E. Prediction of death after withdrawal of life-sustaining treatments. *Critical Care and Resuscitation* 2008; **10**: 278–84.
6. Blackstock MJ, Ray DC. Organ donation after circulatory death: an update. *European Journal of Emergency Medicine* 2014; **21**: 324–9.
7. Taylor LJ, Buffington A, Scalea JR, Fost N, Croes KD, Mezrich JD, Schwarze ML. Harms of unsuccessful donation after circulatory death: an exploratory study. *American Journal of Transplantation* 2018; **18**: 402–9.
8. NHS ODT. *Organ and Tissue Donation and Transplantation Activity Report 2021/22*. NHS Blood and Transplant, 2022.
9. Lewis J, Peltier J, Nelson H, et al. Development of the University of Wisconsin donation after cardiac death evaluation tool. *Progress in Transplantation* 2003; **13**: 265–73.
10. DeVita MA, Brooks MM, Zawistowski C, Rudich S, Daly B, Chaitin E. Donors after cardiac death: validation of identification criteria (DVIC) study for predictors of rapid death. *American Journal of Transplantation* 2008; **8**: 432–41.
11. Davila D, Ciria R, Jassem W, et al. Prediction models of donor arrest and graft utilization in liver transplantation from Maastricht-3 donors after circulatory death. *American Journal of Transplantation* 2012; **12**: 3414–24.
12. de Groot YJ, Lingsma HF, Bakker J, Gommers DA, Steyerberg E, Kompanje EJO. External validation of a prognostic model predicting time of death after withdrawal of life support in neurocritical patients. *Critical Care Medicine* 2012; **40**: 233–8.
13. Rabinstein AA, Yee AH, Mandrekar J, et al. Prediction of potential for organ donation after cardiac death in patients in neurocritical state: a prospective observational study. *Lancet Neurology* 2012; **11**: 414–9.
14. Wind J, Snoeijs MGJ, Brugman CA, et al. Prediction of time of death after withdrawal of life-sustaining treatment in potential donors after cardiac death*. *Critical Care Medicine* 2012; **40**: 766–9.
15. Brieva J, Coleman N, Lacey J, Harrigan P, Lewin TJ, Carter GL. Prediction of death in less than 60 minutes following withdrawal of cardiorespiratory support in ICUs. *Critical Care Medicine* 2013; **41**: 2677–87.

16. Brieva J, Coleman N, Lacey J, Harrigan P, Lewin TJ, Carter GL. Prediction of death in less than 60 minutes after withdrawal of cardiorespiratory support in potential organ donors after circulatory death. *Transplantation* 2014; **98**: 1112–8.

17. He X, Xu G, Liang W, et al. Nomogram for predicting time to death after withdrawal of life-sustaining treatment in patients with devastating neurological injury. *American Journal of Transplantation* 2015; **15**: 2136–42.

18. Okahara S, Snell GI, McDonald M, D'Costa R, Opdam H, Pilcher DV, Levvey B. Improving the predictability of time to death in controlled donation after circulatory death lung donors. *Transplant International* 2021; **34**: 906–15.

19. Kotsopoulos A, Vos P, Witjes M, et al. Prospective multicenter observational cohort study on time to death in potential controlled donation after circulatory death donors-development and external validation of prediction models: the DCD III study. *Transplantation* 2022; **106**: 1844–51. https://doi.org/10.1097/TP.0000000000004106.

20. Kotsopoulos AMM, Boing-Messing F, Jansen NE, Vos P, Abdo WF. External validation of prediction models for time to death in potential donors after circulatory death. *American Journal of Transplantation* 2018; **18**: 890–6.

21. Xu G, Guo Z, Liang W, et al. Prediction of potential for organ donation after circulatory death in neurocritical patients. *Journal of Heart and Lung Transplantation* 2018; **37**: 358–64.

22. Nijhoff MF, Pol RA, Volbeda M, et al. External validation of the DCD-N score and a linear prediction model to identify potential candidates for organ donation after circulatory death: a Nationwide multicenter cohort study. *Transplantation* 2021; **105**: 1311–6.

23. Nicolson C, Burke-Robinson A, Munshi L, Tsanas A, Lone N, Puxty K. Predicting time to asystole following withdrawal of life-sustaining treatment: a systematic review PROSPERO registration. *PROSPERO* 2022 *CRD42022331633*, 2022. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022331633.

24. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine* 2019; **170**: 51–8.

25. Suntharalingam C, Sharples L, Dudley C, Bradley JA, Watson CJE. Time to cardiac death after withdrawal of life-sustaining treatment in potential organ donors. *American Journal of Transplantation* 2009; **9**: 2157–65.

26. Kramer AH, Holliday K, Keenan S, et al. Donation after circulatory determination of death in western Canada: a multicentre study of donor characteristics and critical care practices. *Canadian Journal of Anaesthesia* 2020; **67**: 521–31.

27. Cooke CR, Hotchkin DL, Engelberg RA, Rubinson L, Curtis JR. Predictors of time to death after terminal withdrawal of mechanical ventilation in the ICU. *Chest* 2010; **138**: 289–97.

28. Yee AH, Rabinstein AA, Thapa P, Mandrekar J, Wijdicks EFM. Factors influencing time to death after withdrawal of life support in neurocritical patients. *Neurology* 2010; **74**: 1380–5.

29. Mazer MA, Alligood CM, Wu Q. The infusion of opioids during terminal withdrawal of mechanical ventilation in the medical intensive care unit. *Journal of Pain and Symptom Management* 2011; **42**: 44–51.

30. Huynh TN, Walling AM, Le TX, Kleerup EC, Liu H, Wenger NS. Factors associated with palliative withdrawal of mechanical ventilation and time to death after withdrawal. *Journal of Palliative Medicine* 2013; **16**: 1368–74.

31. Long AC, Muni S, Treece PD, Engelberg RA, Nielsen EL, Fitzpatrick AL, Curtis JR. Time to death after terminal withdrawal of mechanical ventilation: specific respiratory and physiologic parameters may inform physician predictions. *Journal of Palliative Medicine* 2015; **18**: 1040–7.

32. Brudno DS, Boedy RF, Kanto WP. Compliance, alveolar-arterial oxygen difference, and oxygenation index changes in patients managed with extracorporeal membrane oxygenation. *Pediatric Pulmonology* 1990; **9**: 19–23.

33. Munshi L, Dhanani S, Shemie SD, Hornby L, Gore G, Shahin J. Predicting time to death after withdrawal of life-sustaining therapy. *Intensive Care Medicine* 2015; **41**: 1014–28.

34. Dominguez-Gil B, Ascher N, Capron AM, et al. Expanding controlled donation after the circulatory determination of death: statement from an international collaborative. *Intensive Care Medicine* 2021; **47**: 265–81.

Supporting Information

Additional supporting information may be found online via the journal website.

Appendix S1. Embase Classic and Embase search terms.

Appendix S2. Ovid MEDLINE(R) and in-process, in-data-review and other non-indexed citations search terms.

Table S1. Study characteristics.

Table S2. PROBAST data.

Table S3. Top predictors.

Table S4. Model variables.