

Running head: EFFICACY OF SURVIVAL PREDICTION SCORES

Title: Efficacy of outcome prediction of the Respiratory ECMO Survival Prediction Score (RESP) and the Predicting Death for Severe ARDS on VV-ECMO score (PRESERVE) for patients with Acute Respiratory Distress Syndrome on Extracorporeal Membrane Oxygenation.

Authors: Gavin Majithia-Beet MSc^{1,2}, Roozbeh Naemi PhD² and Richard Issitt DClinP^{3,4}

Affiliations:

¹Perfusion Department, Glenfield Hospital, Leicester, UK

²School of Health, Science and Wellbeing, Staffordshire University, Stoke-on-Trent, UK

³Perfusion Department, Great Ormond Street Hospital, London, UK

⁴Institute of Cardiovascular Science, University College London, London, UK

Corresponding Author:

Gavin Majithia-Beet. Department of Perfusion, Glenfield Hospital, University Hospitals of Leicester, Leicester, UK. LE3 9QP. 0116 2563069. gavin.r.majithia-beet@uhl-tr.nhs.uk

EFFICACY OF SURVIVAL PREDICTION SCORES

Abstract

Background: Extracorporeal Membrane Oxygenation (ECMO) therapy for respiratory failure is an increasingly popular modality of support. Patient selection is an important aspect of outcome success. This review assesses the efficacy of the popular prognostic tools RESP and PRESERVE for ECMO patient selection.

Methods: A literature search was performed. 6 publications were found to match the specified selection criteria. These publications were assessed and compared using the area under the receiver operating characteristic (AUROC) curve statistical method to ascertain the discriminatory ability of the models to predict treatment outcome.

Results: Six articles were included in this review from 306 screened, of which all were retrospective cohort studies. Data was generated over a period of 3 to 9 years from 13 referring hospitals. Studies consisted of 467 male and 221 female (30 unknown) participants in total with a high heterogeneity. The PRESERVE prognostic model was found to have a higher AUROC score than the RESP model, however both models were found to be sub-optimal in their discriminatory ability. A high chance of bias was seen across all included studies.

Conclusion: It was the findings of this review, indicated by analysis using the AUROC measures, that the prognostic model PRESERVE performed better than RESP for predicting post ECMO therapy outcomes, for patients presenting with ARDS within their respective validated time frames, i.e. RESP at ICU discharge and PRESERVE at 6 months post ICU discharge. However, it was recognized that comparator groups were small thereby introducing bias into the study. Further prospective, randomized studies would be necessary to effectively assess the utility of these predictive survival scores.

EFFICACY OF SURVIVAL PREDICTION SCORES

Introduction

Acute Respiratory Distress Syndrome (ARDS) accounts for around 10% of Intensive care unit (ICU) admissions carrying a mortality of around 45%¹. This pathology manifests in the form of acute respiratory failure with pulmonary oedema, leading to tissue hypoxia and sometimes hypercapnia. The supportive measures for the management of ARDS include attention to fluid balance with restrictive fluid transfusion strategies, and minimization of sedatives. The use of corticosteroids is also important, and appropriate timely use has been shown to reduce mortality². Finally, the last method of therapy generally used when all other treatments have been exhausted is ECMO.

Since the 1970's, ECMO has been a recognised form of therapeutic, supportive treatment for the failing lungs and/or heart³. When the standard supportive treatment of mechanical ventilation has proven to be insufficient, ECMO has been utilized⁴. This modality of treatment, however, has been used with varying success, it has been apparent that certain cohorts of patients have a better outcome than others and a criterion of contraindications was informally developed⁵. This, along with personal experiences eventually evolved into pragmatic selection criteria which were selectively used by clinicians within their hospitals.

A more empirical, reproducible referral methodology for ECMO was sought by a growing number of clinicians and in 2013 Schmidt et al. proposed the PRESERVE mortality risk score⁶. This group identified risk factors associated with death at six months post ICU discharge for patients with ARDS receiving ECMO therapy. With this data they devised an algorithm whereby patient data could be used to make an informed decision on patient viability; if they would respond well or succumb to their illness by 6 months post ICU discharge.

In 2014, Schmidt et al. suggested another mortality risk score, the RESP prediction score, which predicted the outcome of ARDS patients at ICU discharge⁷. Both models were developed using pre-ECMO variables independently associated with survival at their designated time points using logistic regression. These variables included duration of mechanical ventilation before referral, central nervous system (CNS) dysfunction and age amongst others.

EFFICACY OF SURVIVAL PREDICTION SCORES

No systematic review has yet been carried out on prognostic models for ECMO therapy, yet a definitive opinion on the most effective model would be advantageous to guide future clinical practice. It was the aim of this review to assess the efficacy of the RESP and PRESERVE risk scores in order to determine which was more effective and could then possibly be used to aid clinicians select appropriate candidates for ECMO therapy. Through this judgement, a greater understanding of treatment choice, resource triage and financial efficiency could be used to augment patient care.

Methodology

A comparison of the efficacy of the RESP and PRESERVE mortality risk scores in the clinical setting was carried out by reviewing data on patients who had received ECMO therapy instead of mechanical ventilation alone and comparing the prognostic efficacy of the aforementioned risk models at their defined temporal end points (Resp was survival at ICU discharge and PRESERVE at 6 months post ICU discharge). Studies that did not use the predetermined end points of prognostic models (as specified by the authors that developed the models), were compared using a common endpoint. This is stated in the review and emphasis was made that the model was being used out of context and all results and comparisons postulated.

The review question was developed using the PICO model. The methodology used to identify and procure pertinent articles was based on the 'Preferred Reporting Items for Systematic Reviews' document (PRISMA)⁸.

It should be emphasized that it was the intention of this review to assess the efficacy of the prognostic scores by comparison. The assessment of whether the prognostic scores were themselves intrinsically robust was out of the scope of the current study.

Search strategy

A literature search was performed using two databases, Medline and CINAHL. These were accessed on the Staffordshire University EBSCOhost research platform on the 21/03/21. Medical subject heading

EFFICACY OF SURVIVAL PREDICTION SCORES

(MeSH) terms, Boolean operators and free text searches were applied to a predefined search strategy in order to identify relevant articles. An example of terms used were Extracorporeal membrane oxygenation, ECMO, Extracorporeal Life Support, risk assessment, risk management and mortality. No publication time constraints were applied to the search. Only texts in the English language were searched due to the time available for translation and the concomitant translational bias. All publication types were accepted due to a lack of retrospective studies observed during a scoping search.

Eligibility criteria

In order to be included in this review, all study participants were required to be >17 years of age and the ECMO system must have employed the use of a blood pump. The configuration of the ECMO support must have been veno-venous (VV) and the study cohort needed to be receiving treatment in the ITU setting for ARDS. The mortality prediction models RESP and PRESERVE had to be included; however, other prediction models could be included alongside the models of interest. All other permutations of cannulation for ECMO (i.e., Veno-arterial (VA) and veno-venous-arterial (VVA)) were excluded due to the prediction models in question being postulated solely for VV ECMO patients, and therefore internally validated with this cohort. Post cardiac surgery patients were omitted along with patients that were presenting with any form of COVID-19 respiratory illnesses.

Screening process

The results of the database literature searches were screened by one reviewer (GMB). Initially, screening of the titles and abstracts took place, when any concern or indecision regarding inclusion/exclusion occurred during this process, the full text articles were procured and screened in full. When full text articles were excluded, the reasons were stated on the PRISMA flow diagram (Fig 1).

Data extraction

A data extraction tool was formulated based on the applicable criteria for the narrative synthesis of data. This was developed in conjunction with the Critical Appraisal and Data Extraction for Systematic Reviews

EFFICACY OF SURVIVAL PREDICTION SCORES

of Prediction Modelling Studies (CHARMS) checklist by Moons et al.⁹ Data extraction was completed by one reviewer (GMB).

The extracted data included author, date of extraction, year, study type, number of centers included in study, patient number, primary pathology, survival, death, age, other scores assessed, area under receiver operating characteristics (AUROC) curve, calibration, other tests, missing data and its management, outcomes, conclusions and conflicting interests/funding.

Evaluation of prognostic risk models

The validity of prediction models (diagnostic and prognostic) was assessed by measuring the AUROC curve. RESP and PRESERVE prognostic models were assessed on their predictive accuracy.

Quality assessment

The quality assessment tool used in this study was the Prediction model Risk Of Bias ASessment Tool (PROBAST)¹⁰. This methodology was produced to highlight the risk of bias (ROB) for studies assessing prediction models. This is achieved by applying signaling questions regarding participants, predictors, outcome and analysis to studies in order to assess the overall ROB of a study as well as applicability for concern. Finally, an overall opinion of ROB and applicability is produced by collating this assessment.

Risk of bias across studies

As reviewers during the data extraction process were not blinded there is a chance that confirmation bias could have been present. There was also a possibility of publishing bias occurring, as a limited number of databases were accessed, and these contained only published studies.

EFFICACY OF SURVIVAL PREDICTION SCORES

Results

Synthesis of results

Due to the high level of heterogeneity of studies included in this review because of the disparity between patient characteristics (high clinical diversity), a meta-analysis was not produced. The comparison of studies was not possible because of the differing patient demographic between cohorts, however, because in each study both prognostic prediction models were being evaluated on the same cohort, a narrative synthesis (inter-study comparison) could be made.

Study selection

After the initial screening process 369 articles (Medline n=272 and CINAHL n=97) were identified. After the removal of duplicates 306 remained. Of the 306 records screened 262 were excluded. The remaining 44 full text articles were assessed for eligibility, of which 6 were deemed satisfactory for the inclusion in the review. Of the 38 excluded full text screened articles, 19 did not contain the relevant prognostic models, 11 did not contain both models together for review, 2 utilized pumpless ECMO therapy, 2 provided no comparison between the two models, 3 were correspondence articles and 1 used the prognostic model incorrectly (Table 1).

Overview

Fisser et al. presented the largest cohort of the studies (329) whose data was collected from five referring centers across Europe¹¹. He included over twice as many male participants to females whose body weight was 80kg (no BMI was recorded). His cohort contained a majority of survivors who were younger

EFFICACY OF SURVIVAL PREDICTION SCORES

than the non-survivor group. Lee et al. relied on a relatively small sample (30) from Korea. The ages of this cohort were not stated¹². This group had a relatively low BMI (23.1kg/m²). This was the oldest study, being completed in 2016 but took the least amount of time to collect the data. Kang et al., also demonstrated a relatively smaller BMI of his all Korean study group (22.2 kg/m²) than that of the western groups¹³. Also, with a relatively small sample size, this cohort demonstrated a poor survival rate of 23 patients against 76 that died; however, this study included the oldest patients within the six studies reviewed. The study of Brunet et al., took the longest time to complete, extracting data over nine years from one hospital in France¹⁴. This cohort of 44 patients were the youngest of the studies with a mean age of 44 years, and with over twice as many males this group also had the median BMI of 26.7 kg/m² of the groups. There was an almost equal split outcome of 22 survivors versus 23 deaths. The Hilder et al. study collected data from one German hospital over a period of five years¹⁵. This cohort of 108 patients had the joint largest BMI of 27.8 kg/m², with almost twice as many men to women in the sample population, only one third of patients survived the treatment. The study by Maca et al. had the highest survival rate overall of 65%¹⁶. This study was the most recent from 2021 collecting data from 4 Czech hospitals. Patients notes were accessed from over 8 years creating a data sample size of 111 patients. This study also had the joint largest BMI of 27.8 kg/m² with again, over twice as many males.

In summary, commonalities between the groups were that all were retrospective cohort studies and contained over twice as many male participants as women (467 men compared to 221 women), discluding Lee et al.¹² who did not specify the male/female split in the veno-venous group. All publications were from the last five years. Age ranges of study populations were similar (44-55 years) and within an 11 year range.

Actions on missing data were not specified in any of the studies. Five of the studies reported that all data were available apart from Maca et al.¹⁶ where 18 patients were excluded from the study due to missing data. This omission of patients from this study can introduce sampling bias into the research (Table 2).

Outcome endpoints

EFFICACY OF SURVIVAL PREDICTION SCORES

Four studies used the intended designated reporting period of ICU discharge for the RESP model, but the incorrect period for the PRESERVE model (excluding Maca et al.¹⁶ who reported PRESERVE at 6 months). Brunet et al.¹⁴ gave the incorrect reporting period of 6 months for the RESP model but correct for PRESERVE. Kang et al.¹³ gave a statistical analysis of both ICU discharge and 6 month survival for both models. Hilder et al.¹⁵ did not specify whether their reporting period was ICU discharge or 6 month survival for either model.

Summary of Evidence

Overall, the performance outcome of the prognostic models was at best, sub-optimal. The RESP score had an AUROC curve range of 0.44 to 0.796 (mean= 0.628) in the survival at ICU discharge group (intended use) and 0.60 to 0.69 (mean= 0.645) in the survival at 6 months group (non-intended use), a better performance but the model had not been internally validated for this time period.

The PRESERVE score had an AUROC curve of 0.62-0.69 (mean= 0.657) for the survival at 6 months (intended use) and between 0.60 and 0.69 (mean= 0.645) for the survival at ICU discharge time point (non-intended use).

Hilder et al.¹⁵ showed AUROC curve results for RESP and PRESERVE as 0.645 and 0.593 respectively, however, as no time point for data collection had been stated, it was not possible to give a direct comparison with the other authors findings and as such was not directly compared to other studies in this review.

RESP Model

Lee et al.¹² demonstrated the best functionality of the RESP model with an AUROC curve of 0.7696 (95% confidence level = 0.659-0.897) The Maca et al.¹⁶ study demonstrated the poorest performance of the RESP model of 0.44 (95% confidence level = 0.32-0.55) Fisser et al.¹¹ and Kang et al.¹³ reported an AUROC curve of 0.586 and 0.69 respectively.

PRESERVE Model

EFFICACY OF SURVIVAL PREDICTION SCORES

The study by Brunet et al.¹⁴ showed the best AUROC curve of 0.69 for the PRESERVE model but with wide confidence intervals (95% confidence level = 0.53-0.87).

The poorest performance was described by Maca et al.¹⁶ at 0.62 (95% confidence level = 0.51 - 0.72).

Quality Assessment

The overall risk of bias was high across all studies, with a low applicability by Kang et al.¹³ and Maca et al.¹⁶ (Table 3). All studies scored high for the 'participants' domain due to the methods of enrolment of the patients in the studies, this was due to the studies not being randomized controlled trials. All authors scored low for possible bias in the 'Predictors' domain of the ROB tool due to all prognostic models being assessed in the same way for all participants apart from Hilder et al.¹⁵. Predictor data not being available at the start of the study, namely the time period being utilized for the end point, created high bias. Fisser et al.¹¹ and Lee et al.¹² recorded a high bias for the 'Outcome' domain due to the incorrect choice of end point for the PRESERVE prognostic model, and for Brunet et al.¹⁴ for the RESP model. Hilder et al.¹⁵ does not state an endpoint and as such is considered high also. For the 'Analysis' domain, Lee et al.¹², Brunet et al.¹⁴ and Kang et al.¹³ scored high due to having a sample size of less than 100. Maca et al.¹⁶ scored high due to the mishandling of missing data which resulted in the omission of 18 patients. Hilder et al.¹⁵ scored high due to the relevant performance measures being evaluated inappropriately. The applicability of participants was at a low risk of bias for all studies due to patient samples being as specified in the study. The applicability for predictors was low for all studies apart from Hilder et al.¹⁵ whose timing of predictors in the model do not match the study question.

Discussion

Summary of Evidence

EFFICACY OF SURVIVAL PREDICTION SCORES

In the light of advancing therapeutic options for respiratory failure, the utility of prognostic models to predict the outcome of prospective patients for ECMO therapy is increasing. Empirical rather than experiential methods of patient selection in an evidenced based environment is primarily sought. In our review we assessed the RESP and PRESERVE models of survival after ECMO therapy as these were the most commonly used at the review hospital. It should be noted that according to the Extracorporeal Life Support Organization (ELSO) web site for clinicians, the prognostic scores contained on the web pages of the organization (of which RESP is included) are not recommended for use in determining individual patient management while on ECMO or for patient selection on ECMO. However, Schmidt et al. states in his paper⁷ that “potential roles for the RESP score would include helping clinicians select appropriate candidates for ECMO”. This disparity in guidance may be due to ELSO protecting themselves from prosecution should the models be used and an unfavorable outcome occur, it was certainly the authors intention for the RESP score to have a clinical application for patient assessment.

The function of the models in this review were variable. RESP was seen to provide the best (Lee et al.¹² (0.796)) and worst (Maca et al.¹⁶ (0.44)) ability to discriminate between survival and death in ECMO patients at ICU discharge. The reasons for this were not apparent, the study population for Lee et al.¹² was relatively small (30) with a low BMI (23.1kg/m²) in comparison to those of Maca et al.¹⁶ (111 and 27.8 kg/m² respectively). The PRESERVE score provided consistently sub-optimal discrimination of the binomial end point at the specified 6 month post ICU period.

The PRESERVE score at the ICU discharge time point displayed a better discriminatory ability in its cohorts (458 patients) than at its intended 6 month post ICU discharge time point (251 Patients), the smaller population in the latter could be pertinent with respect to the effect outcome. However, looking at the discriminatory ability of the PRESERVE model over its intended time point of 6 months post ICU discharge, this showed a trend towards the smaller the sample size, the greater the discriminatory ability.

The studies with the better RESP results were by Lee et al.¹² (AUROC curve =0.796) and Kang et al.¹³ (AUROC curve =0.69). These studies were carried out in Korea using an Asian population with the smallest BMI's of all of the studies (23.1 kg/m² and 22.2 kg/m² respectively), this could have contributed to

EFFICACY OF SURVIVAL PREDICTION SCORES

this effect. The studies with the worst RESP results were by Fisser et al.¹¹ and Maca et al.¹⁶, these were the only studies that sourced their data from multiple referring hospitals (5 and 4 respectively).

Relatively small cohorts across the studies, especially with Lee et al.¹² (30) and Brunet et al.¹⁴ (41) could introduce bias and plausibly influence statistical analysis. According to the PROBAST quality assessment tool, validation studies are recommended to have at least 100 patients within their cohort in order to negate possible bias. Comparing these sample sizes to the original studies used to create and validate these prognostic models, the RESP score was originally validated with a population of 2355 patients⁷ and the PRESERVE score 140⁶.

The largest cohort (329) sampled by Fisser et al.¹¹ was collated over a period of eight years from five referring hospitals. Within this large time period, patient treatment and medical techniques could have changed thereby creating a greater heterogeneity between patients, this effect is amplified by collecting patient data from five hospitals. Comparing patients that received ECMO therapy from different units with different treatment plans (no authors stated that patient treatment was standardized across all referring hospitals) would introduce a greater level of heterogeneity. Other authors reporting large durations of data collection were Brunet et al.¹⁴ (9 years), Kang et al.¹³ (8 years) and Maca et al.¹⁶ (8 years).

Both the RESP and PRESERVE score were originally developed retrospectively from a cohort of patients that had undergone ECMO therapy, however, they had both undergone a process of external validation as well as internal¹⁷.

These findings impact upon the viability for the clinical application of these prognostic models. If these models are to be used to decide upon the treatment modality of prospective ECMO patients, they need to be used in conjunction with other methodologies and not as a stand-alone decision tool.

Limitations

The overarching effect of the 6 studies in question being retrospective reviews rather than trials with randomized controlled selection process' is the main concern for data applicability. There is the possibility

EFFICACY OF SURVIVAL PREDICTION SCORES

that as the studies were not prospective, and the data accrued was not intended for research purposes, there could be disparities between the data input and translation.

To reiterate, it was a recognized limitation of the study that the cohorts reported by the authors in question were relatively small thereby introducing the effect of bias. Small sample sizes as well as excessively large have been shown to introduce bias into research¹⁸. Also, the cohorts considered under RESP compared with PRESERVE included a much higher number of patients. These need to be taken into account when the results of this study are interpreted.

Only accepting studies in English may have excluded relevant studies that could have been of value in this review. Due to time constraints, it was deemed acceptable in the given time frame.

Conclusions

It was the findings of this review, based on the analysis with the AUROC statistical method, that the prognostic model PRESERVE performed better than RESP for predicting post ECMO therapy outcomes for patients presenting with ARDS within their respective validated time frames, i.e., RESP at ICU discharge and PRESERVE at 6 months post ICU discharge. Also it was noted that the PRESERVE model provided a greater discriminatory ability at predicting death at hospital discharge than the RESP model, a function to which it was not designed for. This finding came about as a result of some of the aforementioned authors incorrectly assessing the prognostic models. Due to the majority of the studies using the models incorrectly, there was limited data from which conclusions could be drawn, all of the six studies could not be compared on a like for like basis. This superiority in predictive ability in both of these findings was marginal, and both prognostic models should be used with caution and in conjunction with other methodologies if being employed for aiding in the process of patient referral and selection for ECMO therapy. These findings need to be considered in conjunction with the high risk of bias seen throughout all studies in this review.

EFFICACY OF SURVIVAL PREDICTION SCORES

An important conclusion to be drawn from this work is the fact that there have very few studies carried out on the RESP and PRESERVE ECMO prognostic tools to date. It should be noted that the limited studies found that featured in this review were of varying standards, many used RESP and PRESERVE incorrectly and some contained small cohorts of patients creating bias.

It is recommended that future studies in this area of research to fill in the knowledge gaps in a limited area of ECMO treatment should be implemented. Prospective randomized controlled trials need to be carried out to provide studies with less bias, and the ability to capture the specific data required to accurately validate these models. All language publications should be included in the literature search over as many databases as possible. The implications of further research would be that a more empirical approach to patient selection based on scientific fact rather than experiential subjective opinion could be developed giving a more consistent, reproducible effect of ECMO on the treatment of ARDS patients with respiratory failure.

Funding

No funding was sought for this review. All authors declare no competing interests.

References

- [1] Bullen EC, Teijeiro-Paradis R, Fan E. How I Select Which Patients With ARDS Should Be Treated With Venovenous Extracorporeal Membrane Oxygenation. *Chest* 2020; 158: 1036–1045.
- [2] Pierrakos. Acute Respiratory Distress Syndrome: Pathophysiology and Therapeutic Options. *J Clin Med Res* 2012; 4: 7–16.
- [3] Hill JD, O'Brien TG, Murray JJ, et al. Prolonged Extracorporeal Oxygenation for Acute Post-Traumatic Respiratory Failure (Shock-Lung Syndrome): Use of the Bramson Membrane Lung. *N Engl J Med* 1972; 286: 629–634.

EFFICACY OF SURVIVAL PREDICTION SCORES

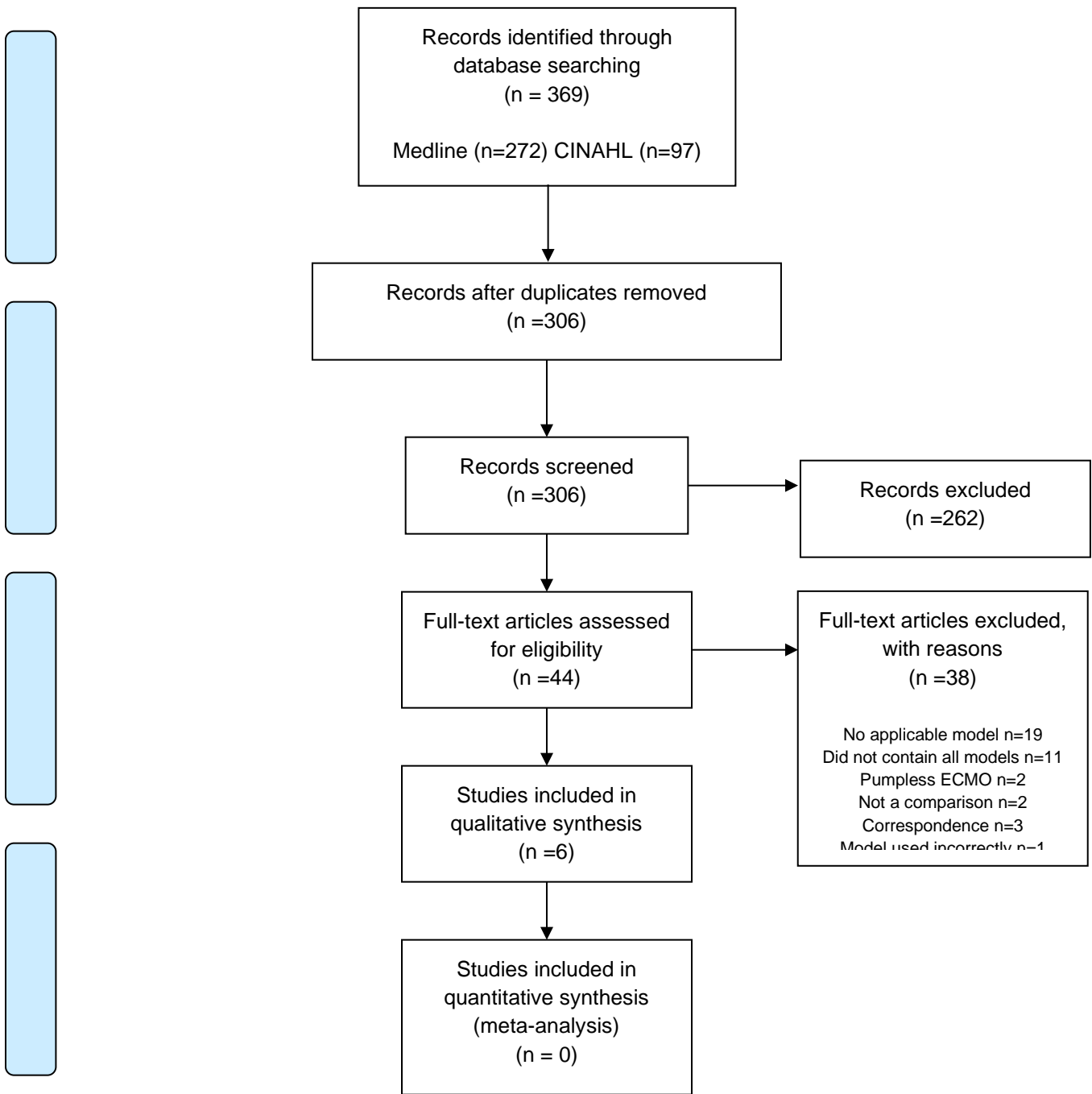
- [4] Mosier JM, Kelsey M, Raz Y, et al. Extracorporeal membrane oxygenation (ECMO) for critically ill adults in the emergency department: History, current applications, and future directions. *Crit Care* 2015; 19: 1–8.
- [5] Makdisi G, Wang IW. Extra Corporeal Membrane Oxygenation (ECMO) review of a lifesaving technology. *J Thorac Dis* 2015; 7: E166–E176.
- [6] Schmidt M, Zogheib E, Rozé H, et al. The PRESERVE mortality risk score and analysis of long-term outcomes after extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. *Intensive Care Med* 2013; 39: 1704–1713.
- [7] Schmidt M, Bailey M, Sheldrake J, et al. Predicting survival after extracorporeal membrane oxygenation for severe acute respiratory failure: The Respiratory Extracorporeal Membrane Oxygenation Survival Prediction (RESP) score. *Am J Respir Crit Care Med* 2014; 189: 1374–1382.
- [8] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*; 6. Epub ahead of print 2009. DOI: 10.1371/journal.pmed.1000097.
- [9] 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*; 11. Epub ahead of print 2014. DOI: 10.1371/journal.pmed.1001744.
- [10] Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. *Ann Intern Med* 2019; 170: W1–W33.
- [11] Fisser C, Rincon-Gutierrez LA, Enger TB, et al. Validation of prognostic scores in extracorporeal life support: A multi-centric retrospective study. *Membranes (Basel)* 2021; 11: 1–12.

EFFICACY OF SURVIVAL PREDICTION SCORES

- [12] Lee S, Yeo HJ, Yoon SH, et al. Validity of Outcome Prediction Scoring Systems in Korean Patients with Severe Adult Respiratory Distress Syndrome Receiving Extracorporeal Membrane Oxygenation Therapy. *J Korean Med Sci* 2016; 31: 932–938.
- [13] Kang H-R, Kim DJ, Lee J, et al. A Comparative Analysis of Survival Prediction Using PRESERVE and RESP Scores. *Ann Thorac Surg* 2017; 104: 797–803.
- [14] Brunet J, Valette X, Buklas D, et al. Predicting Survival After Extracorporeal Membrane Oxygenation for ARDS: An External Validation of RESP and PRESERVE Scores. *Respir Care* 2017; 62: 912–920.
- [15] Hilder M, Herbstreit F, Adamzik M, et al. Comparison of mortality prediction models in acute respiratory distress syndrome undergoing extracorporeal membrane oxygenation and development of a novel prediction score: the PREdiction of Survival on ECMO Therapy-Score (PRESET-Score). *Crit Care* 2017; 21: 301.
- [16] Maca J, Matousek V, Bursa F, et al. Extracorporeal membrane oxygenation survival: external validation of current predictive scoring systems focusing on influenza A etiology. *Artif Organs* 2021; 0–2.
- [17] Klinzing S, Wenger U, Steiger P, et al. External validation of scores proposed for estimation of survival probability of patients with severe adult respiratory distress syndrome undergoing extracorporeal membrane oxygenation therapy: a retrospective study. *Critical Care* 2015; 19: 142-152.
- [18] Faber J, Fonseca L. How sample size influences research outcomes. *Dental Press Journal of Orthodontics* 2014; 19 (4): 27-29

EFFICACY OF SURVIVAL PREDICTION SCORES

Fig 1 PRISMA Flow Chart



EFFICACY OF SURVIVAL PREDICTION SCORES

Table 1. Study Characteristics

Author	Country	Date	Study Type	Duration of study (years)	Study Centres (n)	Sample Size (n)	Age in years	BMI	Sex M/F	Death/age (n/years)	Survival/age (n/years)
Fisser et al. ¹¹	Belgium Italy Sweden Germany	2021	Retrospective Cohort	8	5	329	53 (41-63)	(80)	221/ 108	132/ 56	197/ 50
Lee et al. ¹²	Korea	2016	Retrospective Cohort	3	1	30	NK	23.1	NK	NK	NK
Brunet et al. ¹⁴	France	2017	Retrospective Cohort	9	1	41	44 (26-62)	26.7	29/ 12	23/ 41	22/ 46
Hilder et al. ¹⁵	Germany	2017	Retrospective Cohort	5	1	108	48 (37-55)	27.8	69/ 39	67/ 48	41/ 47
Kang et al. ¹³	Korea	2017	Retrospective Cohort	8	1	99	55 (3872)	22.2	70/ 29	76/NK	23/NK
Maca et al. ¹⁶	Czech Republic	2021	Retrospective Cohort	8	4	111	50 (40-59)	27.8	78/ 33	39 / 55	72 / 48

Ages for the Lee et al. study¹² were given as a combined VV / VA population. In the Kang et al. study¹³ ages of the survivors/non-survivors were not specified. All ages are arithmetic mean averages. For BMI (Body Mass Index), Bracketed values indicate body weight (Kg) in the absence of BMI. n=number, NK= Not Known.

Table 2. RESP/ PRESERVE comparison

Author	RESP		PRESERVE	
	AUROC	95% Confidence Intervals	AUROC	95% Confidence Intervals
Survival at ICU Discharge				
Fisser et al. ¹¹	0.586	0.524-0.649	0.651	0.592-0.710
Lee et al. ¹²	0.796	0.659-0.897	0.80	0.662-0.899
Kang et al. ¹³	0.69	0.58-0.81	0.64	0.51-0.77
Maca et al. ¹⁶	0.44	0.32-0.55		
Hilder et al. ¹⁵	0.645	0.53-0.75	0.593	0.48-0.71
Average	0.631		0.671	
Average (Macca¹⁶,Kang¹³)	0.565			
Survival At 6 Months				
Brunet et al. ¹⁴	0.60	0.41-0.78	0.69	0.53-0.87
Kang et al. ¹³	0.69	0.57-0.81	0.66	0.53-0.79
Maca et al. ¹⁶			0.62	0.51-0.72
Average	0.645		0.657	
Average (Macca¹⁶,Kang¹³)			0.64	

EFFICACY OF SURVIVAL PREDICTION SCORES

Table 3. PROBAST Risk of Bias Assessment Tool

Author	Risk of Bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	ROB	Applicability
Fisser et al. ¹¹	+	-	+	-	-	-	+	+	+
Lee et al. ¹²	+	-	+	+	-	-	+	+	+
Brunet et al. ¹⁴	+	-	+	+	-	-	+	+	+
Hilder et al. ¹⁵	+	-	+	+	-	-	+	+	+
Kang et al. ¹³	+	-	-	+	-	-	-	+	-
Maca et al. ¹⁶	+	-	-	+	-	-	-	+	-

+ indicates high risk of bias or applicability, - indicates low risk of bias or applicability. ROB=Risk Of Bias.