



**QUEEN'S
UNIVERSITY
BELFAST**

Heterogeneity of surrogate outcome measures used in critical care studies: a systematic review

Verghis, R., Blackwood, B., McDowell, C., Toner, P., Hadfield, D., Gordon, A. C., Clarke, M., & McAuley, D. (2023). Heterogeneity of surrogate outcome measures used in critical care studies: a systematic review. *Clinical Trials*, 20(3), 307-318. <https://doi.org/10.1177/17407745231151842>

Published in:
Clinical Trials

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights
Copyright 2023 The Authors.

This is an open access article published under a Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access
This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: <http://go.qub.ac.uk/oa-feedback>

Heterogeneity of surrogate outcome measures used in critical care studies: A systematic review

Clinical Trials
2023, Vol. 20(3) 307–318
© The Author(s) 2023



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17407745231151842
journals.sagepub.com/home/ctj



Rejina Verghis¹ , Bronagh Blackwood¹, Cliona McDowell², Philip Toner¹, Daniel Hadfield³, Anthony C Gordon⁴, Mike Clarke⁵ and Daniel McAuley¹

Abstract

Background: The choice of outcome measure is a critical decision in the design of any clinical trial, but many Phase III clinical trials in critical care fail to detect a difference between the interventions being compared. This may be because the surrogate outcomes used to show beneficial effects in early phase trials (which informed the design of the subsequent Phase III trials) are not valid guides to the differences between the interventions for the main outcomes of the Phase III trials. We undertook a systematic review (1) to generate a list of outcome measures used in critical care trials, (2) to determine the variability in the outcome reporting in the respiratory subgroup and (3) to create a smaller list of potential early phase endpoints in the respiratory subgroup.

Methods: Data related to outcomes were extracted from studies published in the six top-ranked critical care journals between 2010 and 2020. Outcomes were classified into subcategories and categories. A subset of early phase endpoints relevant to the respiratory subgroup was selected for further investigation. The variability of the outcomes and the variability in reporting was investigated.

Results: A total of 6905 references were retrieved and a total of 294 separate outcomes were identified from 58 studies. The outcomes were then classified into 11 categories and 66 subcategories. A subset of 22 outcomes relevant for the respiratory group were identified as potential early phase outcomes. The summary statistics, time points and definitions show the outcomes are analysed and reported in different ways.

Conclusion: The outcome measures were defined, analysed and reported in a variety of ways. This creates difficulties for synthesising data in systematic reviews and planning definitive trials. This review once again highlights an urgent need for standardisation and validation of surrogate outcomes reported in critical care trials. Future work should aim to validate and develop a core outcome set for surrogate outcomes in critical care trials.

Keywords

Clinical trials, critical care, outcomes

Introduction

A clinical endpoint can be defined as a characteristic or variable that reflects on how a patient feels or functions, and how long a patient survives, whereas a surrogate endpoint is a marker intended to substitute for a clinical endpoint that should predict clinical benefit or harm or lack of both.¹ It is often assumed that the surrogate outcomes mimic the clinical endpoint. The validity of a surrogate measure is often hard to prove. One reason for this may be that it is never tested and only a few datasets containing both the surrogate outcome measures and the corresponding clinical endpoint are available for use.

¹The Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, UK

²Northern Ireland Clinical Trials Unit, Belfast, UK

³Critical Care Unit, King's College Hospital NHS Foundation Trust, London, UK

⁴Division of Anaesthetics, Pain Medicine and Intensive Care, Imperial College London, London, UK

⁵Centre of Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, UK

Corresponding author:

Rejina Verghis, The Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast BT9 7BL, UK.
Email: rverghis01@qub.ac.uk

The use of non-validated outcome measures to inform intervention studies, especially confirmatory studies, can cause serious issues. A classic example of this was the Cardiac Arrhythmia Suppression Trial (CAST) I study in 1989 and CAST II in 1992, where the intervention effectively suppressed ventricular arrhythmias (surrogate) but was later shown to increase arrhythmic associated deaths (clinical endpoint).²⁻⁴ In an intensive care unit (ICU)-specific example, the double-blind study of the nitric oxide synthase inhibitor 546C88 showed that the intervention arm significantly improved shock resolution at 72 h (surrogate) in patients with severe sepsis, yet increased the mortality rate (clinical endpoint).⁵⁻⁷

The choice of outcome measure can vary based on the study design and objective. Phase II and III studies have different objectives; Phase II studies aim to provide an initial estimate of the effect size to inform the sample size of a Phase III study and are thus intended to inform the Phase III studies. In Phase II studies, investigators often use a surrogate or a short-term outcome measure to provide an initial estimate of intervention efficacy, whereas Phase III studies more often use a clinical endpoint. For example, the change in Sequential Organ Failure Assessment (SOFA) score, which is a composite of daily physiological measurements, can be considered as an outcome measure for a Phase II study, whereas a Phase III study would use a clinical endpoint such as 60-day mortality or long-term quality of life as the efficacy outcome measure. However, it is not always evident how the effect estimates of Phase II studies, based on surrogate outcome measures, inform the design of Phase III studies, which would use a different outcome measure. For example, how might an effect size based on the change in SOFA score at day 3 reflect changes in 60-day mortality? Many Phase III critical care studies fail to detect a difference between the groups, possibly because the surrogate outcome measures used in early phase studies may not be valid. This is a significant issue as the sample size and resource requirements for Phase III studies are much larger compared with a Phase II study.

Variability of the outcome measures and their definitions is another issue that leads to inconsistencies in research findings and their application. Variability in the outcome measures reported for randomised controlled trials (RCTs) makes any meaningful comparisons between studies difficult. Several authors have previously called for standardisation of outcomes definitions in critical care studies.⁸⁻¹⁰ In 1979, the World Health Organization (WHO) published a handbook on results reporting in cancer studies, which is considered the earliest attempt to standardise outcome reporting in healthcare research.¹¹ The Outcome Measures in Rheumatoid Arthritis Clinical Trials Group (OMERACT) was the first group to formally recognise the variability in outcome selection and

reporting in Rheumatology studies. The OMERACT network was initiated in 1992 and consensus conferences run biennially.¹² The recommendations are data-driven and are updated by relevant working groups. The OMERACT model was followed by other medical fields including critical care. The Core Outcome Measures in Effectiveness Trial (COMET) initiative launched in 2010 leads the creation of Core Outcome Sets to improve standardisation and reporting in effectiveness studies.⁹ The International Forum for Acute Care Trialists (InFACT) collaboration established in 1989 drives outcomes research in critical care studies.¹³ The National Heart, Lung and Blood Institute and the Society of Critical Care Medicine had highlighted a need to gain consensus on a standard set of long-term outcomes for post-ICU discharge studies.¹⁴

The aim of this systematic review was to determine different outcomes reported in critical care trials, to determine the variability in outcome reporting in the respiratory subgroup, and to create a smaller list of potential early phase endpoints for critical care trials in the respiratory subgroup.

Methods

This section details the systematic review methodology and the details of data extraction and data synthesis.

Types of studies

Critical care RCTs involving adult patients were included in this review. All design types were included. Studies published in six journals in the critical care category with the highest impact factor at the time were included in this review. This restriction was made on the assumption that the studies would provide a good representation of outcomes reported in critical care trials. All interventions, pharmacological, non-pharmacological and medical device were included in this review, based on the assumption that the studies aimed to improve the clinical and efficacy outcomes of the patients and these outcomes were similar across several types of interventions.

Type of participants

Studies were included if the participants were adult patients in intensive care units. Healthy volunteer studies, paediatric studies, end of life studies and transplant studies were excluded.

Search methods

Studies were those published in the *American Journal of Respiratory and Critical Care Medicine*, *Chest*, *Critical Care Medicine*, *Critical Care Medicine*, *Intensive Care Medicine* and *Lancet Respiratory Medicine*, between the years 2010

and 2020. The search used MeSH headings, keywords and variants for 'intensive care unit' or 'critical care' combined (using the Boolean operator AND) with search strings to identify RCTs. The review was conducted according to the protocol published on the PROSPERO website (http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015017607).¹⁵

Studies investigating interventions for Corona Virus Disease 2019 (COVID-19) were not included in this review. A separate search was conducted on <https://clinicaltrials.gov/> database, on Jan 2022, to identify the outcomes as per the trial registration for COVID-19 studies. Search terms used were COVID-19, COVID and corona virus. The outcomes identified in this review were compared with COVID study outcomes.

Study selection

One reviewer conducted the initial search and screening. The study title, author names, abstract and journal names were extracted. Two reviewers examined the title and abstract of studies identified by the search. Full text of the studies deemed potentially suitable were retrieved and read to confirm eligibility. All studies published in 2010 which reported on surrogate outcome measures were considered for data extraction, and 10% of studies from 2014 and 2015. A further search was conducted in July 2020 and 10% of articles published in the past year were also extracted to see whether there was any significant change in the outcomes reported in these more recent years. The eligible articles were listed in an Excel sheet and 10% of the articles were randomly selected from the list.

Data extraction

Details on all outcomes reported by the treatment arm were extracted. Baseline characteristics were not extracted. Details on the trial registration, outcome measures, definitions, time points and statistics were extracted.

The specific measurement variable, which corresponds to the data collected directly from trial participants (e.g. SOFA score); the participant-level analysis metric, which corresponds to the format of the outcome data that will be used from each trial participant for analysis (e.g. change from baseline, time to event); the method of aggregation, which refers to the summary measure format for each study group (e.g. mean, proportion); and the specific measurement time point of interest for analysis were extracted.

Data cleaning and analysis

All outcomes except mortality and quality of life outcomes were considered as a surrogate outcome for this

review. Outcomes related to safety and compliance were excluded since safety and compliance were not considered as outcomes that would give an initial estimate of the drug efficacy. Similarly, outcomes that were not specified in the methods section of the article were also excluded. If a study reported on SOFA score at days 1, 3 and 7, it was counted as one outcome. If a study reported the absolute SOFA score and change in SOFA (e.g. from baseline to day 7), this was counted as two outcomes.

At first, outcomes were arbitrarily categorised into body organ systems, biomarkers, disease severity score and resource use, and then into subcategories. Composite outcomes were placed into one category based on the most relevant component in the outcome. For example, ventilator free days which consist of components mortality and mechanical ventilation were classified under mechanical ventilation, because mortality outcomes were not included in this review and mechanical ventilation was the most relevant component. A smaller list of outcomes relevant to the respiratory subgroup was then considered for further analysis. The clinical relevance of the outcome was determined by one of the authors who is an ICU clinician and a leading critical care researcher. All data were entered into Microsoft Excel for analysis. Outcomes were tabulated to understand the patterns and variability. Pivot charts, pivot tables and sun charts were used to report the results. Counts and percentages were used to summarise the results.

Results

A total of 6905 references were retrieved and 4046 were excluded at the initial screening as the inclusion criteria were not met. Titles and abstracts of 2859 studies were reviewed. Full text of 465 studies was retrieved. Data were extracted from 58 studies.¹⁶⁻⁷⁴ The flow chart in Figure 1 shows the study selection process.

Study characteristics

Studies included pharmacological 33 (56.9%), non-pharmacological 16 (27.6%) and medical device 9 (15.5%) interventions. Table 1 summarises the intervention type and patient condition. Nineteen (32.8%) were infection-related studies and 16 (27.6%) involved respiratory illness.

Outcome categories and subcategories

The outcome measures, analysis metrics, time points and aggregation methods from the studies were identifiable for most of the outcomes. However, these parameters varied from trial to trial. Outcome measures were re-extracted for five studies by the second

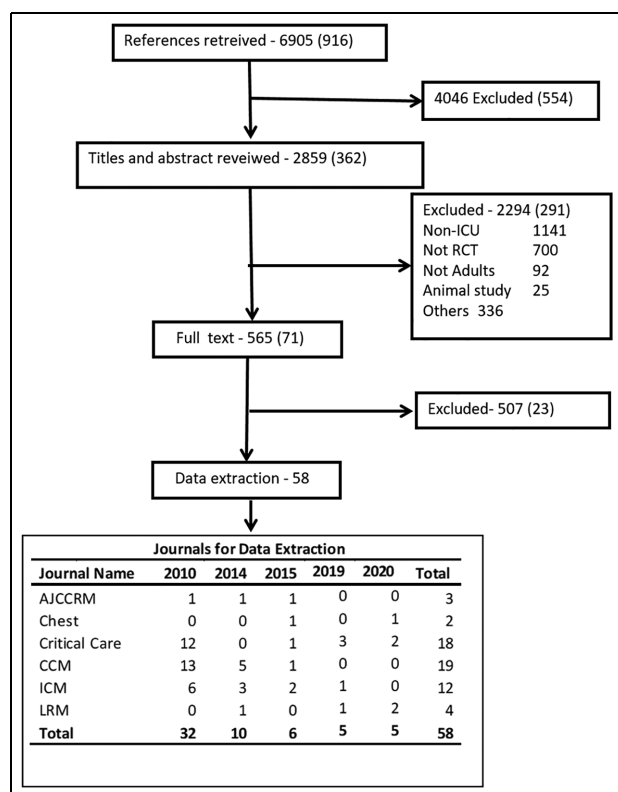


Figure 1. Study flowchart.

AJRCCM: American Journal of Respiratory and Critical Care Medicine; CCM: Critical Care Medicine; ICM: Intensive Care Medicine; LRM: Lancet Respiratory Medicine. Numbers for 2010 are given in brackets.

Table 1. Study characteristics.

	N
Total number of articles	58
Intervention	
Pharmacological	33 (56.9%)
Non-pharmacological	16 (27.6%)
Medical device	9 (15.5%)
Condition	
Infections and infestations	19 (32.8%)
Respiratory, thoracic, and mediastinal disorders	16 (27.6%)
Nervous system disorders	9 (15.5%)
General critical illness	4 (6.9%)
Cardiac disorders	3 (5.2%)
Renal and urinary disorders	2 (3.4%)
Metabolism and nutrition disorders	2 (3.4%)
Gastrointestinal disorders	2 (3.4%)
Surgical and medical procedures	1 (1.7%)

reviewer. There was 100% agreement after resolving the differences.

A total of 294 separate outcomes were identified from 58 studies. The 294 outcome measures were grouped into 11 categories and 66 subcategories; 33 (57%) studies reported on resource use outcome and 30 (52%) studies reported on cardiovascular outcomes.

There were 50 (17%) cardiovascular outcomes, 41 (18%) respiratory outcomes and 46 (16%) infection outcomes. Table 2 shows the outcome categories, sub-categories, number of outcomes identified and number of studies reporting the outcomes. Appendix C in the supplementary material elaborates on the outcomes and categories in Table 2.

Length of ICU stay was reported 25 times by 24 studies, which indicates that multiple definitions were used to report on ICU stay. For example, if the ICU stay was reported for all patients and for survivors in one study, these were considered as two different outcomes. Length of stay outcomes, mechanical ventilation outcomes and SOFA were the most popular outcomes. The others were physiological outcomes, and the majority of these were cardiovascular outcomes.

Potential early phase endpoints for respiratory subgroup in ICU

Based on the National Health Service (NHS) digital data published in 2017 and 2020, approximately 30% of ICU patients will require advanced respiratory support and approximately 80% will require some form of respiratory support,⁷⁵⁻⁷⁷ indicating that the respiratory subgroup is one of the most burdened requiring better research and development. Hence a smaller list of 22 outcomes out of the 294 outcomes was chosen, based on the clinical relevance for the subgroup. Table 3 shows the outcome, number of articles (%) reporting the outcome, analysis metric used and aggregation method for these 22 outcomes.

Organ dysfunction outcomes: Duration of mechanical ventilation, organ failure free days and ventilator free days were the organ dysfunction outcomes identified as the potential early phase endpoints. Outcomes related to mechanical ventilation were frequently reported and consisted of the most variable definitions. Mechanical ventilation outcomes definitions included the duration of mechanical ventilation, rate of mechanical ventilation, mechanical ventilation free days, number of ventilated days, time to successful extubation and ventilator free days. Mechanical ventilation outcomes were reported using mean (standard deviation), mean (standard error), median (range), median (interquartile range), median (range), Kaplan–Meier estimate and median (95% confidence interval) Kaplan–Meier estimate.

Length of stay outcomes: Length of stay in ICU and hospital and ICU free days were reported using mean (standard deviation), mean (standard error), median (range), median (inter-quartile range), and median (95% confidence interval) Kaplan–Meier estimate. Length of stay outcome definition was reported for all patients and for survivors' only.

Table 2. Summary table on outcome classification.

Category	Number of subcategories identified	Number of outcomes identified	Number of studies reporting the outcomes
Biomarker	2	20	14
Blood and lymphatic system	2	16	10
Cardiovascular system	6	50	30
Hepatobiliary system	1	5	8
Infection	13	46	22
Metabolism and nutrition system	6	18	10
Nervous system	12	33	13
Renal and urinary system	4	19	21
Resource	5	12	33
Respiratory, thoracic and mediastinal system	11	54	41
Severity of disease	5	19	21
Grand total	66	294	58

Table 3. Potential early phase outcomes, analysis metric, aggregation method for the respiratory subgroup.

Outcome	N ^a (%)	Analysis metric	Aggregation method	Time point
Duration of mechanical ventilation	20 (34.5%)	Duration of mechanical ventilation – all patients Duration of mechanical ventilation – survivors Duration of mechanical ventilation (invasive and non-invasive) Duration of mechanical ventilation until weaning Days off mechanical ventilation Intubation free days Time before first weaning attempt Time to weaning from mechanical ventilation Weaning duration	Mean \pm SD, median (IQR), mean \pm SE, median (range), Kaplan–Meier estimate	Extubation, during study period, day 28
Duration of ICU stay	27 (46.6%)	Duration of ICU stay – all patients Duration of ICU stay – survivors only Duration of level 2 ICU stay Duration of level 3 ICU stay	Mean \pm SD, median (IQR), Mean \pm SE, Median (range)	At ICU discharge, during study period, days 21 and 28
Duration of hospital stay	20 (34.5%)	Duration of hospital stay–all patients Duration of hospital stay–survivors	Mean \pm SD, Median (IQR), mean \pm SE, median (range)	At hospital discharge, during study period, day 21
ICU-free days	3 (5.2%)	ICU free days	Median (IQR), mean \pm SE	28 days
Organ failure–free days	3 (5.2%)	Organ failure free days	n/N, %, n (%)	During study period, 60 days
Ventilator-free days	13 (22.4%)	Ventilator free day	Mean, SD, median (IQR), mean \pm SE	21 and 28 days
SOFA score	14 (24.1%)	SOFA (absolute, change, maximum) Non-hepatic SOFA (absolute, change) SOFA corrected (absolute, change)	Mean \pm SD, mean (95% CI), median (IQR), mean \pm SE	0, 0–6, 8, 0–8 and 9–72 h, baseline, days 0, 1, 3, 5, 7, 10, 14, 21 and 28
MOD score	1 (1.7%)	Absolute value	Text stating ‘no significant difference’	
Lung injury score	1 (1.7%)	Absolute value	Mean \pm SD	Baseline, days 2 and 4

(continued)

Table 3. Continued

Outcome	N ^a (%)	Analysis metric	Aggregation method	Time point
Oxygenation index	2 (3.4%)	Absolute value	Mean ± SD, Mean ± SE	Baseline, days 0, 1, 2, 3, 4, 5 and 7
PF ratio	8 (13.8%)	Absolute value, time spend with PaO ₂ : FiO ₂ < 200 mm Hg	Mean ± SD, Mean ± SE	Baseline, 1, 3, 6, 12, 24, 36 and 48 h, days 0, 1, 2, 3, 4, 5, 6 and 7
Platelets	4 (6.9%)	Absolute value, change score	N, mean, SD, min, max, mean ± SD, median (IQR)	Baseline, 1, 12, 24, 48, 72 and 120 h, days 1, 2 and 6
CRP	1 (1.7%)	Absolute value	Mean ± SD, range	Days 1, 2, 3, 6, 18 and 14
Urine output	1 (1.7%)	Fluid intake, fluid output, fluid balance	Mean ± SD, median (IQR), adjusted mean (95% CI)	Baseline, 8, 0–8 and 9–72 h
Creatinine	4 (6.9%)	Absolute value, change values	Mean, SD, min, max	1, 2, 24, 48, 72 and 120 h
IL-10	4 (6.9%)	Absolute value, change value	Mean ± SD, median (IQR), mean ± SE, range	Baseline, 48–72 h, days 1 and 6
IL-1B	5 (8.6%)	Absolute value, change value	Mean ± SD, mean ± SE, range	0, 8, 0–8 and 9–72 h
IL-6	9 (15.5%)	Absolute value, change value	AUC, mean ± SD, median (IQR), mean ± SE, range	0, 2, 6, 8, 24, 36, 48, 48–72, 72 and 96.5 h, days 1, 2, 3, 4, 6, 7, 8 and 14
IL-8	5 (8.6%)	Absolute value, change value	Mean ± SD, median (IQR), range	0, 8, 24, 48–72 and 96.5 h, days 1, 2 and 4
sRAGE	2 (3.4%)	Absolute value	AUC, mean ± SD, mean ± SE, AUC, sensitivity, specificity	–5, 5 and 30 min, 1, 4 and 6 h, baseline, days 1, 2 and 4
TNFα	7 (12.1%)	Absolute value, change value	Mean ± SD, median, mean ± SE, range	0, 8 and 25 h, baseline, days 1, 2, 4 and 6
CRs static	1 (1.7%)	Absolute value	Mean ± SD	Baseline days 2 and 4

SD: standard deviation; IQR: interquartile range; ICU: Intensive care unit; PF ratio: PaCO₂/FiO₂; MOD: Multiple Organ Dysfunction; CRP: C-reactive protein; SOFA: sequential organ failure assessment; CI: confidence interval; SE: standard error; AUC: area under the curve.

^aNumber of articles and percentage (%).

Disease severity scores: Severity scores in Table 3 include SOFA score, MOD score and Lung Injury Score. All three scores are calculated from physiology scores that are routinely collected in the ICU. SOFA was defined in numerous ways such as non-hepatic SOFA and SOFA corrected values which exclude the neurology component. Organ failure was defined using SOFA using different cut-offs such as a change in SOFA > 2 and a SOFA score > 6. SOFA and MOD scores indicate multiple organ failures.

Physiology outcomes: Physiology outcomes were classified into routinely collected data and biomarkers. Looking at the frequency of the outcomes reported, a total of 136/305 (55%) were short-term physiology outcomes, which are routinely collected in the ICUs. Other outcomes are the biomarkers.

Figure 2 proposes the classification of outcomes that should be reported on an early phase critical care trial. The idea is that all early phase critical care trials report on an organ dysfunction outcome, a length of stay outcome, a disease severity score and physiology outcomes. The physiological outcome should be the one related to the underlying condition. For example, Oxygenation

Index and PaO₂/FiO₂ ratio (PF ratio) is a physiological outcome associated with ARDS. An early phase study on ARDS patient population should report on the duration of ventilation, length of ICU stay, length of hospital stay, SOFA score, PF ratio and Oxygenation Index. Similarly, a cardiology study may report on a heart dysfunction measure, length of stay, SOFA score, heart rate and a relevant biomarker.

This review did not include COVID studies, and two letters related to COVID studies were identified during article screening. The search on clinicaltrials.gov database for COVID studies identified 191 studies. A total of 2379 outcomes were specified in the database for 191 studies and the number of outcomes per study ranged from 1 to 74. Mortality-related outcomes were specified 227 times, free days scores were specified 306 times and 160 studies specified ventilator free day as an outcome. COVID-19 ordinal clinical progression outcome scale was another outcome frequently specified; however the days and numbers in the scale varied. Other frequently reported outcomes were duration of mechanical ventilation, ICU/hospital length of stay and SOFA.

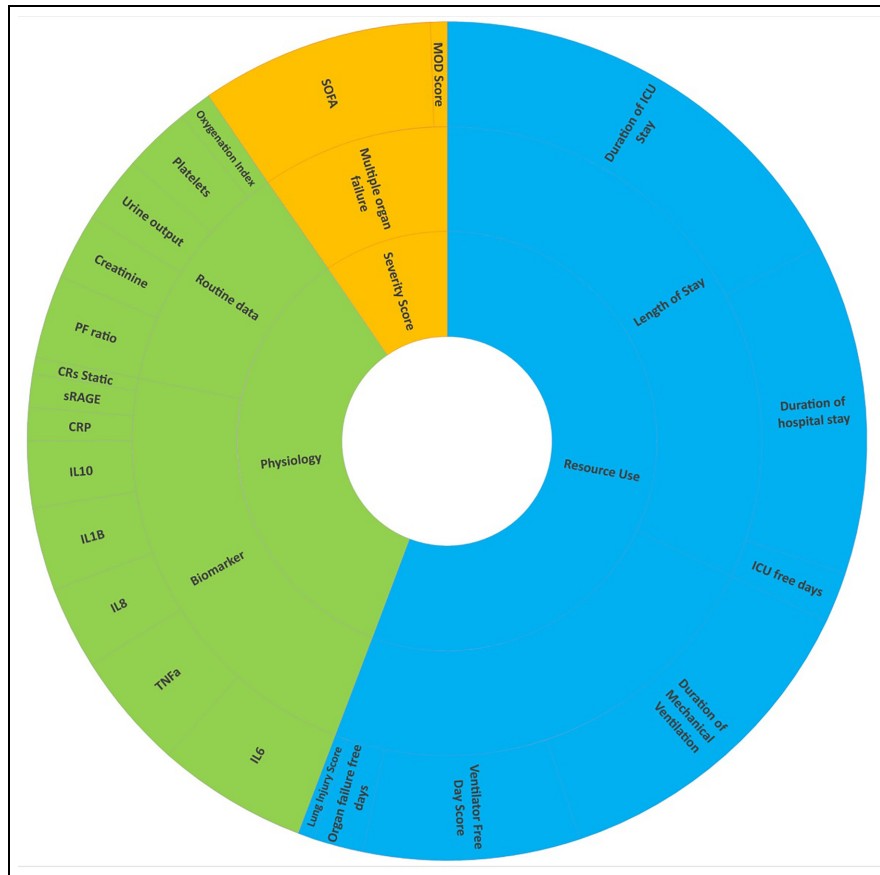


Figure 2. Potential early phase endpoints in respiratory subgroup and classifications (outer circle shows the outcomes, mid-circle represents the subcategory and the inner circle represents the category).

Conclusion

This qualitative review looked at surrogate measures and their variability in critical care studies. This review identified about 294 different outcomes. These outcomes were defined and reported clearly within most of the studies, however, their definitions and reporting varied from study to study. Review of trial registration details of the COVID study showed that variability of outcomes and outcome reporting is an issue in COVID studies as well. This makes the comparison between the studies difficult, which shows that there is an urgent need to standardise the outcomes.

Most of the outcomes found in the review can be broadly classified into organ dysfunction outcomes, length of stay outcomes, disease severity scores, routine physiology data and biomarkers. This classification is similar to those proposed in the work of Dodd et al.,⁷⁸ except for adverse events, given the exclusion of outcomes related to adverse events from this review. An estimate of the change in disease severity can be expected to provide an estimate for 'life impact' outcomes such as quality of life. A subset of 22 outcomes relevant for the respiratory subgroup was selected for further research. Mechanical ventilation, ventilator free

days and organ failure free days indicate organ dysfunction. Mechanical ventilation is a lifesaving intervention in the ICU and ventilation requirements have direct effects on resource use and a direct and attributable effect on mortality. Blackwood et al.⁷⁹ have previously demonstrated the variation in the measurement and reporting of mechanical ventilation outcomes. Furthermore, Contentin et al.⁸⁰ reviewed 128 reports on adult ICU studies and identified 13 different definitions of ventilator free days. The variability in the definition and analysis of mechanical ventilation outcomes was evident in this review as well. The variability in ventilation approaches varies from ICU to ICU and this influences variability even further. Mechanical ventilation outcomes, ventilator free days and organ failure free days indicate organ dysfunction, which increases the risk of death. Organ dysfunction would also impact the length of stay of the patient in the hospital and ICU.

Length of stay outcomes are relevant to clinicians and patients. ICU/hospital length of stay outcomes are clear indicators of resource use. ICU/hospital free days were recommended as endpoints of Phase II studies by the Australia and New Zealand Intensive Care Society

(ANZICS) group.⁸¹ The definition of the length of stay can vary based on whether it is an intervention or medical device study. Intervention studies usually measure the length of stay from the day of randomisation or patient enrolment to the study, while a medical device or an observational study is more likely to measure the length of stay from ICU admission. A point to remember is that a reduction in the duration of mechanical ventilation and length of stay can also be due to patients dying early. Hence combining survivors and non-survivors when reporting the results can be misleading and these outcomes should be reported separately by survival status.

Physiology outcomes were subcategorised into routinely collected outcomes and biomarkers. In Table 3, 13 out of 22 outcomes can be categorised as physiology outcomes. Physiology outcomes such as PF ratio and Oxygenation Index are primary and secondary outcome measures in several ICU studies.^{82–85} However, the magnitude of the association between these variables and mortality is not clear. In the ICU, multiple physiology outcomes can be combined to generate severity scores such as SOFA, which are associated with mortality. Recent developments in data analysis techniques allow the use of these measurements to make predictions.

Many authors have previously reported on the variability of outcome measures in healthcare research. The results of this review once again highlight issues of variability among surrogate outcomes and inconsistency of outcome reporting in critical care studies. Systematic reviews and meta-analyses compare and combine the evidence from various research studies carried out in a field or on an intervention. Heterogeneity in outcome reporting makes a comparison between the studies exceedingly difficult and time-consuming. This is not an effective use of resources and time invested in healthcare research. A Core Outcome Set is the minimum set of outcomes that should be collected and reported. Studies can still collect and report other relevant outcomes and need not be restricted to the outcomes in the core outcome set. The development of a core outcome set and standardisation of these outcomes will reduce the impact of outcome variability.

There are a few limitations to note in this review. First is that the review included all adult ICU studies, except healthy volunteer, end of life and transplant studies. The study was based on the work of the ANZICS group⁸¹ and was not narrowed down to a specific group of patients, even though conditions requiring ICU care are indiscriminate and heterogeneous. They can cover the entire spectrum of medicine in aetiology from trauma to cardiovascular to psychiatric. These patients often require complex interventions. All these factors might have had an impact on the variability of outcomes. The second limitation is that the study was restricted to those published in the top

six journals in critical care over a brief period. Data extraction was limited to the published reports of the studies and protocols were not checked for definitions of outcomes. The impact of this is thought to be minimal because the review was able to identify the different outcomes used in critical care and the variable definitions used which was the purpose of the study. A third limitation was the exclusion of safety outcomes. However, we emphasise that safety reporting is crucial in all phases of clinical studies. A fourth limitation relates to the outcomes being skewed towards respiratory studies. Three of the six high impact journals had a focus on respiratory system, and this could have led to this. The set of outcomes identified can be broadly classified as organ failure outcomes, length of stay outcomes, routinely collected physiology outcome and biomarkers. Similar sets of outcomes can be identified for other patient groups for further testing. A fifth limitation concerns the selection of the articles; all articles from 2010 were included in the review and 10% from other years. The 10% was randomly selected, and this could have potentially excluded a few important outcomes.

Acknowledgement

The authors acknowledge NICTU for funding the project. They also acknowledge Gordon Rubinfeld, Sunnybrook Health Sciences, Toronto; Prof. Ranjith Lall, Warwick University and Dr Murali Shyamsundar, Queens University Belfast, for their independent review.

Author contribution

D.F.K. and R.V. conceived the study. R.V. conducted the analysis and drafted the manuscript. R.V. and D.H. conducted the article selection and outcome extraction. R.V., P.T. and D.H. did the outcome classifications. D.F.K. did the outcome selection in the respiratory category. B.B., C.M., D.F.K. and M.C. made substantial contribution in relation to the reviewing and supervising. A.C.G. made intellectual contribution to the manuscript. All authors have read and approved the manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: R.V. received Doctoral Research Fellowship grant from Northern Ireland Clinical Trials Unit (NICTU) for the project. A.C.G. is supported by an NIHR Research Professorship (grant no. RP-2015-06018) and the NIHR Imperial Biomedical Research Centre.

Consent for publishing

Not applicable.


Data statement

Not Applicable.

Ethics approval

Not applicable.

ORCID iD

Rejina Verghis  <https://orcid.org/0000-0002-8322-213X>

Supplemental material

Supplemental material for this article is available online.

References

- Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001; 69(3): 89–95.
- Ruskin JN. The cardiac arrhythmia suppression trial (CAST). *The New England J Med* 1989; 321 (6): 386–388.
- Greene HL, Roden DM, Katz RJ, et al. The cardiac arrhythmia suppression trial: first CAST... then CAST-II. *J Am Coll Cardiol* 1992; 19(5): 894–898.
- Epstein AE, Bigger JT Jr, Wyse DG, et al. Events in the Cardiac Arrhythmia Suppression Trial (CAST): mortality in the entire population enrolled. *J Am Coll Cardiol* 1991; 18(1): 14–19.
- Bakker J, Grover R, McLuckie A, et al. Administration of the nitric oxide synthase inhibitor NG-methyl-L-arginine hydrochloride (546C88) by intravenous infusion for up to 72 hours can promote the resolution of shock in patients with severe sepsis: results of a randomized, double-blind, placebo-controlled multicenter study (study no. 144-002). *Crit Care Med* 2004; 32(1): 1–12.
- López A, Lorente JA, Steingrub J, et al. Multiple-center, randomized, placebo-controlled, double-blind study of the nitric oxide synthase inhibitor 546C88: effect on survival in patients with septic shock. *Crit Care Med* 2004; 32(1): 21–30.
- Watson D, Grover R, Anzueto A, et al. Cardiovascular effects of the nitric oxide synthase inhibitor NG-methyl-L-arginine hydrochloride (546C88) in patients with septic shock: results of a randomized, double-blind, placebo-controlled multicenter study (study no. 144-002). *Crit Care Med* 2004; 32(1): 13–20.
- Clarke M and Williamson PR. Core outcome sets and systematic reviews. *Systematic Reviews* 2016; 5: 1–4.
- Williamson P and Clarke M. Editorial: The COMET (Core Outcome Measures in Effectiveness Trials) initiative: its role in improving Cochrane review. *Cochr Datab Syst Rev* 2012; 5: 1–3.
- Kirkham JJ, Gorst S, Altman DG, et al. Core Outcome Set–STAndards for Reporting: the COS-STAR statement. *Plos Med* 2016; 13(10): e1002148.
- World Health Organization. *WHO handbook for reporting results of cancer treatment*. Geneva: World Health Organization, 1979.
- Tugwell P, Boers M, Brooks P, et al. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007; 8: 1–6.
- Collaboration IGHN. InFACT: a global critical care research response to H1N1. *Lancet (London, England)* 2010; 375: 11–13.
- Dinglas VD, Faraone LN and Needham DM. Understanding patient-important outcomes after critical illness: a synthesis of recent qualitative, empirical, and consensus-related studies. *Curr Opin Crit Care* 2018; 24(5): 401–409.
- Verghis R, Blackwood B, McDowell C, et al. Surrogate outcomes reported in critical care trials, 2016. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015017607
- Acosta-Escribano J, Fernández-Vivas M, Grau Carmona T, et al. Gastric versus transpyloric feeding in severe traumatic brain injury: a prospective, randomized trial. *Intensive Care Med* 2010; 36(9): 1532–1539.
- Albert M, Williamson D, Muscedere J, et al. Candida in the respiratory tract secretions of critically ill patients and the impact of antifungal treatment: a randomized placebo controlled pilot trial (CANTREAT study). *Intensive Care Med* 2014; 40(9): 1313–1322.
- Barbosa VM, Miles EA, Calhau C, et al. Effects of a fish oil containing lipid emulsion on plasma phospholipid fatty acids, inflammatory markers, and clinical outcomes in septic patients: a randomized, controlled clinical trial. *Crit Care* 2010; 14(1): R5–R11.
- Barraud D, Blard C, Hein F, et al. Probiotics in the critically ill patient: a double blind, randomized, placebo-controlled trial. *Intensive Care Med* 2010; 36(9): 1540–1547.
- Boerma EC, Koopmans M, Konijn A, et al. Effects of nitroglycerin on sublingual microcirculatory blood flow in patients with severe sepsis/septic shock after a strict resuscitation protocol: a double-blind randomized placebo controlled trial. *Crit Care Med* 2010; 38(1): 93–100.
- Casey E, Lane A, Kuriakose D, et al. Bolus remifentanyl for chest drain removal in ICU: a randomized double-blind comparison of three modes of analgesia in post-cardiac surgical patients. *Intensive Care Med* 2010; 36(8): 1380–1385.
- Constantin J-M, Futier E, Cherprenet A-L, et al. A recruitment maneuver increases oxygenation after intubation of hypoxemic intensive care unit patients: a randomized controlled study. *Crit Care* 2010; 14: R76.
- Deane AM, Chapman MJ, Fraser RJ, et al. Effects of exogenous glucagon-like peptide-1 on gastric emptying and glucose absorption in the critically ill: relationship to glycemia. *Crit Care Med* 2010; 38(5): 1261–1269.
- Determann RM, Royackers A, Wolthuis EK, et al. Ventilation with lower tidal volumes as compared with conventional tidal volumes for patients without acute lung injury: a preventive randomized controlled trial. *Crit Care* 2010; 14: R1.
- Devlin JW, Roberts RJ, Fong JJ, et al. Efficacy and safety of quetiapine in critically ill patients with delirium: a prospective, multicenter, randomized, double-blind, placebo-controlled pilot study. *Crit Care Med* 2010; 38(2): 419–427.

26. Dixon B, Schultz MJ, Smith R, et al. Nebulized heparin is associated with fewer days of mechanical ventilation in critically ill patients: a randomized controlled trial. *Crit Care* 2010; 14(5): R180.
27. Galiatsatos P, Gibson BR, Rabiee A, et al. The glucoregulatory benefits of glucagon-like peptide-1 (7-36) amide infusion during intensive insulin therapy in critically ill surgical patients: a pilot study. *Crit Care Med* 2014; 42(3): 638–645.
28. Gattas DJ, Rajbhandari D, Bradford C, et al. A randomized controlled trial of regional citrate versus regional heparin anticoagulation for continuous renal replacement therapy in critically ill adults. *Crit Care Med* 2015; 43(8): 1622–1629.
29. Girard R, Baboi L, Ayzac L, et al. The impact of patient positioning on pressure ulcers in patients with severe ARDS: results from a multicentre randomised controlled trial on prone positioning. *Intensive Care Med* 2014; 40(3): 397–403.
30. Girard TD, Pandharipande PP, Carson SS, et al. Feasibility, efficacy, and safety of antipsychotics for intensive care unit delirium: the MIND randomized, placebo-controlled trial. *Crit Care Med* 2010; 38(2): 428–437.
31. Gordon AC, Mason AJ, Perkins GD, et al. The interaction of vasopressin and corticosteroids in septic shock: a pilot randomized controlled trial. *Crit Care Med* 2014; 42(6): 1325–1333.
32. Jabaudon M, Hamroun N, Roszyk L, et al. Effects of a recruitment maneuver on plasma levels of soluble RAGE in patients with diffuse acute respiratory distress syndrome: a prospective randomized crossover study. *Intensive Care Med* 2015; 41(5): 846–855.
33. Jansen TC, van Bommel J, Schoonderbeek FJ, et al. Early lactate-guided therapy in intensive care unit patients: a multicenter, open-label, randomized controlled trial. *Am J Resp Crit Care Med* 2010; 182: 752–761.
34. Jhanji S, Vivian-Smith A, Lucena-Amaro S, et al. Haemodynamic optimisation improves tissue microvascular flow and oxygenation after major surgery: a randomised controlled trial. *Crit Care* 2010; 14(4): R151.
35. Jones C, Backman C, Capuzzo M, et al. Intensive care diaries reduce new onset post traumatic stress disorder following critical illness: a randomised, controlled trial. *Crit Care* 2010; 14(5): R168.
36. Jung YJ, Koh Y, Hong SB, et al. Effect of vancomycin plus rifampicin in the treatment of nosocomial methicillin-resistant *Staphylococcus aureus* pneumonia. *Crit Care Med* 2010; 38(1): 175–180.
37. Köhnlein T, Windisch W, Köhler D, et al. Non-invasive positive pressure ventilation for the treatment of severe stable chronic obstructive pulmonary disease: a prospective, multicentre, randomised, controlled clinical trial. *Lancet Respir Med* 2014; 2(9): 698–705.
38. Kalfon P, Giraudeau B, Ichai C, et al. Tight computerized versus conventional glucose control in the ICU: a randomized controlled trial. *Inten Care Med* 2014; 40: 171–181.
39. Kirakli C, Naz I, Ediboglu O, et al. A randomized controlled trial comparing the ventilation duration between adaptive support ventilation and pressure assist/control ventilation in medical patients in the ICU. *Chest* 2015; 147(6): 1503–1509.
40. Magder S, Potter BJ, Varennes BD, et al. Fluids after cardiac surgery: a pilot study of the use of colloids versus crystalloids. *Crit Care Med* 2010; 38(11): 2117–2124.
41. Maggiore SM, Richard JC, Abroug F, et al. A multicenter, randomized trial of noninvasive ventilation with helium-oxygen mixture in exacerbations of chronic obstructive lung disease. *Crit Care Med* 2010; 38(1): 145–151.
42. Morelli A, Donati A, Ertmer C, et al. Levosimendan for resuscitating the microcirculation in patients with septic shock: a randomized controlled study. *Crit Care* 2010; 14(6): R232.
43. Morris PE, Promes JT, Guntupalli KK, et al. A multicenter, randomized, double-blind, parallel, placebo-controlled trial to evaluate the efficacy, safety, and pharmacokinetics of intravenous ibuprofen for the treatment of fever in critically ill and non-critically ill adults. *Blood* 2010; 9: 29.
44. Mueller SW, Preslaski CR, Kiser TH, et al. A randomized, double-blind, placebo-controlled dose range study of dexmedetomidine as adjunctive therapy for alcohol withdrawal. *Crit Care Med* 2014; 42(5): 1131–1139.
45. Pérez-Bárcena J, Crespi C, Regueiro V, et al. Lack of effect of glutamine administration to boost the innate immune system response in trauma patients in the intensive care unit. *Crit Care* 2010; 14(6): R233.
46. Parienti JJ, Megarbane B, Fischer MO, et al. Catheter dysfunction and dialysis performance according to vascular access among 736 critically ill adults requiring renal replacement therapy: a randomized controlled study. *Crit Care Med* 2010; 38(4): 1118–1125.
47. Payen DM, Guilhot J, Launey Y, et al. Early use of polymyxin B hemoperfusion in patients with septic shock due to peritonitis: a multicenter randomized control trial. *Intensive Care Med* 2015; 41(6): 975–984.
48. Pettilä V, Kyhälä L, Kylänpää ML, et al. APCAP – activated protein C in acute pancreatitis: a double-blind randomized human pilot trial. *Crit Care* 2010; 14(4): R139.
49. Prondzinsky R, Lemm H, Swyter M, et al. Intra-aortic balloon counterpulsation in patients with acute myocardial infarction complicated by cardiogenic shock: the prospective, randomized IABP SHOCK Trial for attenuation of multiorgan dysfunction syndrome. *Crit Care Med* 2010; 38(1): 152–160.
50. Rice TW, Wheeler AP, Bernard GR, et al. A randomized, double-blind, placebo-controlled trial of TAK-242 for the treatment of severe sepsis. *Crit Care Med* 2010; 38(8): 1685–1694.
51. Richard J-C, Bayle F, Bourdin G, et al. Preload dependence indices to titrate volume expansion during septic shock: a randomized controlled trial. *Crit Care* 2015; 19: 5.
52. Robinson S, Zincuk A, Strøm T, et al. Enoxaparin, effective dosage for intensive care patients: double-blinded, randomised clinical trial. *Crit Care* 2010; 14(2): R41.
53. Routsis C, Gerovasili V, Vasileiadis I, et al. Electrical muscle stimulation prevents critical illness polyneuromyopathy: a randomized parallel intervention trial. *Crit Care* 2010; 14: R74.
54. Seder CW, Stockdale W, Hale L, et al. Nasal bridling decreases feeding tube dislodgment and may increase caloric intake in the surgical intensive care unit: a randomized, controlled trial. *Crit Care Med* 2010; 38(3): 797–801.

55. Seguin P, Laviolle B, Dahyot-Fizelier C, et al. Effect of oropharyngeal povidone-iodine preventive oral care on ventilator-associated pneumonia in severely brain-injured or cerebral hemorrhage patients: a multicenter, randomized controlled trial. *Crit Care Med* 2014; 42(1): 1–8.
56. Torgersen C, Dünser MW, Wenzel V, et al. Comparing two different arginine vasopressin doses in advanced vasodilatory shock: a randomized, controlled, open-label trial. *Intensive Care Med* 2010; 36(1): 57–65.
57. Trof RJ, Sukul SP, Twisk JW, et al. Greater cardiac response of colloid than saline fluid loading in septic and non-septic critically ill patients with clinical hypovolaemia. *Intensive Care Med* 2010; 36(4): 697–701.
58. Trzeciak S, Glaspey LJ, Dellinger RP, et al. Randomized controlled trial of inhaled nitric oxide for the treatment of microcirculatory dysfunction in patients with sepsis. *Crit Care Med* 2014; 42(12): 2482–2492.
59. Walz JM, Avelar RL, Longtine KJ, et al. Anti-infective external coating of central venous catheters: a randomized, noninferiority trial comparing 5-fluorouracil with chlorhexidine/silver sulfadiazine in preventing catheter colonization. *Crit Care Med* 2010; 38(11): 2095–2102.
60. Westermaier T, Stetter C, Vince GH, et al. Prophylactic intravenous magnesium sulfate for treatment of aneurysmal subarachnoid hemorrhage: a randomized, placebo-controlled, clinical study. *Crit Care Med* 2010; 38(5): 1284–1290.
61. Zhu B, Li Z, Jiang L, et al. Effect of a quality improvement program on weaning from mechanical ventilation: a cluster randomized trial. *Intensive Care Med* 2015; 41(10): 1781–1790.
62. Maggiore SM, Idone FA, Vaschetto R, et al. Nasal high-flow versus Venturi mask oxygen therapy after extubation. Effects on oxygenation, comfort, and clinical outcome. *Am J Resp Crit Care Med* 2014; 190: 282–288.
63. Vaschetto R, Longhini F, Persona P, et al. Early extubation followed by immediate noninvasive ventilation vs. standard extubation in hypoxemic patients: a randomized clinical trial. *Intensive Care Med* 2019; 45(1): 62–71.
64. Sitbon O, Bosch J, Cottrel E, et al. Macitentan for the treatment of portopulmonary hypertension (PORTICO): a multicentre, randomised, double-blind, placebo-controlled, phase 4 trial. *Lancet Respir Med* 2019; 7(7): 594–604.
65. He H, Sun B, Liang L, et al. A multicenter RCT of noninvasive ventilation in pneumonia-induced early mild acute respiratory distress syndrome. *Crit Care* 2019; 23: 1–13.
66. Constantin JM, Jabaudon M, Lefrant JY, et al. Personalised mechanical ventilation tailored to lung morphology versus low positive end-expiratory pressure for patients with acute respiratory distress syndrome in France (the LIVE study): a multicentre, single-blind, randomised controlled trial. *Lancet Respir Med* 2019; 7(10): 870–880.
67. Mistraretti G, Umbrello M, Salini S, et al. Enteral versus intravenous approach for the sedation of critically ill patients: a randomized and controlled trial. *Crit Care* 2019; 23: 1–10.
68. Hajjar LA, Zambolim C, Belletti A, et al. Vasopressin versus norepinephrine for the management of septic shock in cancer patients: the VANCS II randomized clinical trial. *Crit Care Med* 2019; 47(12): 1743–1750.
69. Miller J, Bruen C, Schnaus M, et al. Auxora versus standard of care for the treatment of severe or critical COVID-19 pneumonia: results from a randomized controlled trial. *Crit Care* 2020; 24: 1–9.
70. Borges I, Carneiro R, Bergo R, et al. Duration of antibiotic therapy in critically ill patients: a randomized controlled trial of a clinical and C-reactive protein-based protocol versus an evidence-based best practice strategy without biomarkers. *Crit Care* 2020; 24: 1–11.
71. Hadfield DJ, Rose L, Reid F, et al. Neurally adjusted ventilatory assist versus pressure support ventilation: a randomized controlled feasibility trial performed in patients at risk of prolonged mechanical ventilation. *Crit Care* 2020; 24: 1–10.
72. Hellyer TP, McAuley DF, Walsh TS, et al. Biomarker-guided antibiotic stewardship in suspected ventilator-associated pneumonia (VAPrapid2): a randomised controlled trial and process evaluation. *Lancet Respir Med* 2020; 8(2): 182–191.
73. Villar J, Ferrando C, Martínez D, et al. Dexamethasone treatment for the acute respiratory distress syndrome: a multicentre, randomised controlled trial. *Lancet Respir Med* 2020; 8(3): 267–276.
74. Chang P, Liao Y, Guan J, et al. Combined treatment with hydrocortisone, vitamin C, and thiamine for sepsis and septic shock: a randomized controlled trial. *Chest* 2020; 158(1): 174–182.
75. Esteban A, Frutos-Vivar F, Muriel A, et al. Evolution of mortality over time in patients receiving mechanical ventilation. *Am J Resp Crit Care Med* 2013; 188: 220–230.
76. Digital N. Hospital adult critical care activity 2015-16, 2017, <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-adult-critical-care-activity/2015-16#:~:text=Patients%20aged%2050%20years%20and,of%20all%20critical%20care%20records>
77. Digital N. Hospital admitted patient care activity 2019-20, 2020, <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2019-20>
78. Dodd S, Clarke M, Becker L, et al. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol* 2018; 96: 84–92.
79. Blackwood B, Clarke M, McAuley DF, et al. How outcomes are defined in clinical trials of mechanically ventilated adults and children. *Am J Resp Crit Care Med* 2014; 189: 886–893.
80. Contentin L, Ehrmann S and Giraudeau B. Heterogeneity in the definition of mechanical ventilation duration and ventilator-free days. *Am J Resp Crit Care Med* 2014; 189: 998–1002.
81. Young P, Hodgson C, Dulhunty J, et al. End points for phase II trials in intensive care: recommendations from the Australian and New Zealand Clinical Trials Group consensus panel meeting. *Crit Care Resusc* 2012; 14: 211–215.
82. Toner P, O’Cane C, McNamee J, et al. Aspirin as a treatment for acute respiratory distress syndrome – a

- multi-centre, randomised, double-blind, placebo-controlled trial (STAR): study protocol. *Crit Care Horiz* 2018; 1: 1–7.
83. McAuley DF, Laffey JG, O’Kane CM, et al. Simvastatin in the acute respiratory distress syndrome. *N Engl J Med* 2014; 371: 1695–1703.
84. Shyamsundar M, McAuley DF, Ingram RJ, et al. Keratinocyte growth factor promotes epithelial survival and resolution in a human model of lung injury. *Am J Resp Crit Care Med* 2014; 189: 1520–1529.
85. Craig TR, Duffy MJ, Shyamsundar M, et al. A randomized clinical trial of hydroxymethylglutaryl-coenzyme a reductase inhibition for acute lung injury (the HARP study). *Am J Resp Crit Care Med* 2011; 183: 620–626.