Received: 10 August 2022 | Accepted: 17 August 2022

DOI: 10.1111/aas.14142

RESEARCH ARTICLE

acta Anaesthesiologica Scandinavica

Causal inference for planning randomised critical care trials: Protocol for a scoping review

Benjamin Skov Kaas-Hansen[1,2] | Anders Granholm[1] | Carl Thomas Anthon[1] | Maj-Brit Nørregaard Kjær[1] | Praleene Sivapalan[1] | Mathias Maagaard[3] | Olav Lilleholt Schjørring[4,5] | Steen Kåre Fagerberg[4] | Karen Louise Ellekjær[1] | Jesper Mølgaard[6] | Claus Thorn Ekstrøm[2] | Morten Hylander Møller[1] | Anders Perner[1]

[1]Department of Intensive Care, Copenhagen University Hospital, Copenhagen, Denmark

[2]Section for Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[3]Department of Anaesthesiology, Centre for Anaesthesiological Research, Zealand University Hospital Køge, Køge, Denmark

[4]Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Aalborg, Denmark

[5]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

[6]Department of Anesthesiology, Centre for Cancer and Organ Dysfunction, Copenhagen University Hospital, Copenhagen, Denmark

Correspondence
Benjamin Skov Kaas-Hansen, Department of Intensive Care, Copenhagen University Hospital, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen E, Denmark.
Email: bkaa0011@regionh.dk; epiben@hey.com

Funding information
Dagmar Marshalls Fond; Grosserer Jakob Ehrenreich og Hustru Grete Ehrenreichs Fond; Sygeforsikringen "danmark"

Abstract

Background: Randomised clinical trials in critical care are prone to inconclusiveness owing, in part, to undue optimism about effect sizes and suboptimal accounting for heterogeneous treatment effects. Planned predictive enrichment based on secondary critical care data (often very rich with respect to both data types and temporal granularity) and causal inference methods may help overcome these challenges, but no overview exists about their use to this end.

Methods: We will conduct a scoping review to assess the extent and nature of the use of causal inference from secondary data for planned predictive enrichment of randomised clinical trials in critical care. We will systematically search 10 general and specialty journals for reports published on or after 1 January 2018, of randomised clinical trials enrolling adult critically ill patients. We will collect trial metadata (e.g., recruitment period and phase) and, when available, information pertaining to the focus of the review (predictive enrichment based on causal inference estimates from secondary data): causal inference methods, estimation techniques and software used; types of patient populations; data provenance, types and models; and the availability of the data (public or not). The results will be reported in a descriptive manner.

Discussion: The outlined scoping review aims to assess the use of causal inference methods and secondary data for planned predictive enrichment in randomised critical care trials. This will help guide methodological improvements to increase the utility, and facilitate the use, of causal inference estimates when planning such trials in the future.

1 | INTRODUCTION

Randomised clinical trials (RCTs) in critical care are prone to inconclusiveness owing, in part, to undue optimism about effect sizes[1,2] or suboptimal accounting for heterogeneous treatment effects (HTEs).[3,4] To some extent, these key challenges may be addressed by prudent enrichment, that is, an informed approach to maximising trial efficiency that can be planned (while designing the trial)[5] or adaptive (when the trial is underway).[6,7] Both types can be prognostic and predictive: the former seeks to identify participants more likely to

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Acta Anaesthesiologica Scandinavica* published by John Wiley & Sons Ltd on behalf of Acta Anaesthesiologica Scandinavica Foundation.

experience the outcome under study, the latter those more likely to respond to the intervention.[5] Accordingly, planned enrichment relies on data collected prior to trial initiation whereas adaptive enrichment uses data collected during the trial.

The intensive care unit (ICU) setting enjoys rich and heterogeneous data along several axes[8,9]: data type (e.g., clinical observations, medicine, biochemistry); structuredness (e.g., tabular, free text); and temporal granularity, from stationary (e.g., date of birth, biological sex) over high-frequency (e.g., ventilator modality, continuous medicine infusion) to waveforms (e.g., continuous blood pressure, electrocardiograms). This richness potentially makes observational ICU data a strong basis for planned predictive (and prognostic) enrichment in RCTs.

Indeed, multiple methods exist for estimating causal effects in observational data, including propensity scores,[10,11] sufficient covariate sets derived from directed acyclic graphs (DAGs),[12,13] and causal discovery.[14–17] Despite their existence and firm methodological bearing[18,19]—with some specifically targeting HTEs[20]—little is known about their use in RCTs in critical care.

Thus, we here lay out the details of a scoping review elucidating the extent and nature of the use of causal inference from secondary data for planned predictive enrichment of RCTs in critical care.

2 | MATERIALS AND METHODS

This protocol adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P)[21] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR)[22]; filled-in checklists are provided in Appendix S1.

2.1 | Research questions

This scoping review will be undertaken in the context of the Intensive Care Platform Trial (INCEPT; www.incept.dk) research programme, to inform, for example, the crafting of the analytic infrastructure to undertake planned predictive enrichment in randomised critical care trials.[5] We aim to answer five questions:

1. How many recent RCTs in critical care used causal inference methods for planned predictive enrichment?
2. What causal inference methodologies have been used for planned predictive enrichment in these RCTs?
3. What kinds of data were used?
4. What data models were used?
5. To what extent were publicly available data used?

2.2 | Search strategy

We will use and update the search and data extraction of a recent scoping review on patient-important outcomes in critically ill patients[23,24]: RCTs in critical care whose results were published online or in-print on or after 1 January 2018 in 10 general and specialised journals (American Journal of Respiratory and Critical Care Medicine, BMJ, Chest, Critial Care, Critical Care Medicine, Intensive Care Medicine, JAMA, The Lancet, The Lancet Respiratory Medicine, New England Journal of Medicine). The start date of the query was arbitrary but chosen to yield a sufficiently large convenience sample of (conceivably well-conducted) RCTs. The full search string (including the Cochrane Collaboration's highly sensitive filter for randomised clinical trials[25]) is detailed in Appendix S1.

2.3 | Eligibility

We will include RCTs enrolling mainly patients admitted to ICUs or similar high-dependency units, defined as patients receiving critical care-specific treatment (e.g., mechanical ventilation, continuous use of vasopressors/inotropes, continuous renal replacement therapy, extracorporeal membrane oxygenation). Quasi- and cluster-randomised trials will be included. For RCTs enrolling also non-ICU patients, we will base our in- or exclusion on the descriptive summary statistics to establish whether the majority (>50%) of patients could reasonably be assumed critically ill, as done previously.[24]

We will exclude trials with patient-level cross-over designs, trials in which the interventions primarily happen outside the ICU (e.g., perioperative interventions in patients subsequently admitted to the ICU), and trials whose interventions happen after discharge from the ICU (even if participants are randomised while in the ICU).[24]

2.4 | Screening and selection

When updating the search of the above-mentioned scoping review,[24] we will use Covidence (www.covidence.org) for title/abstract and full-text screening. All articles will be assessed independently in duplicate, resolving disagreements by consensus and involving a third assessor if need be.

2.5 | Data sources, extraction and synthesis

For any given included trial we will have up to four data sources: the principal reports, the statistical analysis plans, protocols (published or appended to the principal reports), and responses from the corresponding author (only if information is not otherwise available; see below). We will use a custom online form for data entry; screenshots of the form will be included in the supplement to the final scoping review. Data extraction will start with a pilot phase in which two assessors extract data from 10 trials, to test the data extraction form and make pertinent changes before full extraction starts. As for the screening phase, all data will be extracted independently in duplicate, resolving disagreements by consensus and involving a third assessor if need be.

For each included RCT we will extract:

- Study acronym/first author of principal report
- Digital object identifier of principal report
- Journal of principal report
- Year of publication of principal report
- Intervention type(s): drug, medical device, management (as previously defined[24,26,27])
- Recruitment period
- Number of enrolled participants
- Number of centres
- Whether the trial was stopped early and, if so, the reason for termination (e.g., futility or superiority) and whether this was according to a predefined stopping rule or not (i.e., post-hoc decision)
- Phase, using also public trial registers such as ClinicalTrials.gov when necessary (e.g., phase-III registration trial and phase-IV post-marketing trial)
- Whether the RCT was restricted to COVID-19 patients

For each causal inference method we will extract:

- Method name(s) (e.g., generalised linear model [although not strictly as causal inference method in itself], propensity-score based, stratification-score based, hierarchical modelling, counterfactuals-as-missing, pre-specified DAG, causal discovery, causal structure learning). We will build the list of methods on a running basis during data extraction for speedier and harmonised data entry.
- Estimation technique(s) (e.g., frequentist, fully Bayesian)
- Software (language and library/ies)

For each data set used for predictive enrichment we will extract:

- Country/ies of origin
- Number of subjects
- Patient population (e.g., in-hospital internal medicine, primary sector and previously conducted RCT(s))
- Provenance (e.g., clinical observations, ICU apparatus data and national registers)
- Data types (e.g., biochemistry, medication and diagnoses)
- Data model (e.g., the Observational Medical Outcomes Partnership (OMOP) common data model,[28] the Fast Healthcare Interoperability Resources (FIHR) framework,[29] idiosyncratic)
- Vocabularies/ontologies for idiosyncratic data models (e.g., the International Classification of Disease version 10 [ICD-10], the Anatomical Therapeutical Chemical [ATC] classification system and the Systematised Nomenclature of Medicine [SNOMED])
- Whether the data are private or publicly available (e.g., the Medical Information Mart of Intensive Care (MIMIC,[30]) and the eICU Collaborative Research Database[31])

If additional variables are identified in the pilot phase, we will expand the lists above accordingly and report this in the scoping review.

We consult the main text of the principal report, its statistical analysis plan and/or the protocol (if published) to obtain information about the use of causal inference from secondary data for predictive enrichment; we consider secondary data to be "data generated for a purpose different from the research activity for which they are used" (chap. 23 in Rothman et al.[13]). We will contact the corresponding authors of RCTs for which neither of those three sources mentions the use of causal inference methods for planned predictive enrichment. After at least 2 weeks we will send a second enquiry if no response was received to the first; if no response is received 2 weeks later, we will assume no causal inference methods were used for planned predictive enrichment in the given trials.

In line with the scoping review's exploratory nature and due to the multitude of causal inference methods available, we will not specify exactly what constitutes a *causal inference method* a priori (see examples above, however) but instead decide on a case-by-case basis during extraction, including the final decisions in the published scoping review.

We will summarise categorical data with absolute and relative frequencies, and numeric data with medians and inter-quartile ranges (very skewed and sparse numeric variables will be binned and reported as ordinal data). Falling outside the scope of this review type, we will assess neither risk of bias nor certainty of evidence[22]; further, there will be no null-hypothesis significance tests. We will quantify and report inter-rater agreement during data extraction with Cohen's kappa.[32]

2.6 | Ethics and reporting

Ethics approval is not required as we will use data available in the public domain. Regardless of their nature, the results will be published in an international peer-reviewed scientific journal, duly reporting and justifying any deviations from this protocol. In a similar vein, decisions made and conventions developed during screening will be reported in the supplement.

3 | DISCUSSION

In the outlined scoping review we aim to describe the use of causal inference in critical care trials. This may elicit applications of these methods to bolster future RCTs in critical care through planned predictive enrichment. The results may also elucidate shortcomings of existing methods to guide methodological development in the area to render the methods applicable in both critical care trials specifically and clinical trials overall.

The principal strengths of the outlined scoping review include its adhering to standards in the field of scoping reviews[21,22]; the prior publication of this protocol and use of trusted software with logged screening history; and the inclusion of also traditional regression analyses even though such are often misspecified or might even fail to estimate the intended treatment effects.[33]

The proposed review will, however, have some limitations. First, we only include RCTs from which results were published in or after 2018. Albeit arbitrary, we expect this timespan to yield enough RCTs (the scoping review whose search we will update included 167 trials[24]) to sufficiently reflect the state of contemporary RCTs in critical care. Further, many advances in the area of causal inference have occurred in the past decade or so. Second, we restrict the search to four general and six specialised journals. These do, however, serve as major outlets of randomised critical care trials, and so, although our search strategy may miss some, we expect to identify most well-conducted RCTs in critical care. Third, we focus on predictive enrichment and omit RCTs employing prognostic enrichment. These two enrichments schemes, however, are distinct with different methodological considerations,[5] and so we found it prudent to keep them apart to enable a concerted discussion of our results. Because the data sources will be similar for prognostic enrichment, much of the work to be done in this scoping review could be reused in a separate future scoping review of planned prognostic enrichment.

4 | CONCLUSIONS

The outlined scoping review aims to provide an overview of the extent and nature of the use of causal inference estimates from secondary data for planned predictive enrichment of RCTs in critical care. This may inform and facilitate the use of causal inference in future RCTs in critical care, for example, through methodological improvements.

AUTHOR CONTRIBUTIONS

Conceptualisation: Benjamin Skov Kaas-Hansen. Study design: all authors. Writing, original draft: Benjamin Skov Kaas-Hansen. Writing, critical review: all authors. Guarantor: Anders Perner. All authors approved the final version of the protocol. Funders played no role in designing this study.

ACKNOWLEDGMENTS

The outlined study will be conducted as part of the Intensive Care Platform Trial (INCEPT; www.incept.dk) research programme, which has received funding by Sygeforsikringen "danmark", Grosserer Jakob Ehrenreich og Hustru Grete Ehrenreichs Fond and Dagmar Marshalls Fond.

CONFLICT OF INTEREST

All authors declare no personal conflicts of interest. The Department of Intensive Care at Rigshospitalet: has received funding for other research projects from the Novo Nordisk Foundation, Pfizer, Ferring and Fresenius Kabi, and conducts contract research for AM-Pharma.

ORCID

Benjamin Skov Kaas-Hansen https://orcid.org/0000-0003-1023-0371
Anders Granholm https://orcid.org/0000-0001-5799-7655
Carl Thomas Anthon https://orcid.org/0000-0001-7740-700X
Maj-Brit Nørregaard Kjær https://orcid.org/0000-0002-6536-0504
Praleene Sivapalan https://orcid.org/0000-0002-0442-0032
Mathias Maagaard https://orcid.org/0000-0002-9037-7295
Olav Lilleholt Schjørring https://orcid.org/0000-0002-7749-6003
Steen Kåre Fagerberg https://orcid.org/0000-0001-5846-0155
Karen Louise Ellekjær https://orcid.org/0000-0001-9724-9504
Jesper Mølgaard https://orcid.org/0000-0003-2062-9570
Claus Thorn Ekstrøm https://orcid.org/0000-0003-1191-373X
Morten Hylander Møller https://orcid.org/0000-0002-6378-9673
Anders Perner https://orcid.org/0000-0002-4668-0123

REFERENCES

1. Ridgeon EE, Bellomo R, Aberegg SK, et al. Effect sizes in ongoing randomized controlled critical care trials. *Crit Care*. 2017;21:132. doi:10.1186/s13054-017-1726-x
2. Aberegg SK, Richards DR, O'Brien JM. Delta inflation: a bias in the design of randomized controlled trials in critical care medicine. *Crit Care*. 2010;14:R77. doi:10.1186/cc8990
3. Abrams D, Montesi SB, Moore SKL, et al. Powering bias and clinically important treatment effects in randomized trials of critical illness. *Crit Care Med*. 2020;48:1710-1719. doi:10.1097/CCM.0000000000004568
4. Harhay MO, Wagner J, Ratcliffe SJ, et al. Outcomes and statistical power in adult critical care randomized trials. *Am J Respir Crit Care Med*. 2014;189:1469-1478. doi:10.1164/rccm.201401-0056CP
5. FDA. Enrichment strategies for clinical trials to support determination of effectiveness of human drugs and biological products: guidance for industry. 2019. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/enrichment-strategies-clinical-trials-support-approval-human-drugs-and-biological-products.
6. Angus DC, Alexander BM, Berry S, et al. Adaptive platform trials: definition, design, conduct and reporting considerations. *Nat Rev Drug Discov*. 2019;18:797-807. doi:10.1038/s41573-019-0034-3
7. Simon N, Simon R. Adaptive enrichment designs for clinical trials. *Biostatistics (Oxford, England)*. 2013;14:613-625. doi:10.1093/biostatistics/kxt010
8. Thorsen-Meyer H. *Prognostication in Intensive Care Using Machine Learning*. University of Copenhagen; 2020
9. Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care. *Chest*. 2018;154:1239-1248. doi:10.1016/j.chest.2018.04.037
10. Hernán MA, Robins JM. *Causal Inference: What if*. Chapman & Hall/CRC; 2020.
11. Walker A, Patrick AR, Lauer MS, et al. A tool for assessing the feasibility of comparative effectiveness research. *J Comp Eff Res*. 2013;3:11-20. doi:10.2147/CER.S40357
12. Pearl J. *Causality: Models, Reasoning and Inference*. 2nd ed. Cambridge University Press; 2009.
13. Rothman KJ, Lash TL, Greenland S. *Modern Epidemiology*. 3rd ed. Lippincott Williams & Wilkins; 2012.
14. Peters J, Bühlmann P, Meinshausen N. Causal inference using invariant prediction: identification and confidence intervals. *arXiv (unpublished)*. 2015. doi:10.48550/ARXIV.1501.01332
15. Petersen AH, Osler M, Ekstrøm CT. Data-driven model building for life-course epidemiology. *Am J Epidemiol*. 2021;190:1898-1907. doi:10.1093/aje/kwab087
16. Spirtes P, Glymour C. An algorithm for fast recovery of sparse causal graphs. *Soc Sci Comp Rev*. 1991;9:62-72. doi:10.1177/089443939100900106
17. Wu J, Horeweg N, de Bruyn M, et al. Automated causal inference in application to randomized controlled clinical trials. *arXiv (unpublished)*. 2022. doi:10.48550/ARXIV.2201.05773
18. Blakely T, Lynch J, Simons K, Bentley R, Rose S. Reflection on modern methods: when worlds collide-prediction, machine learning and causal

inference. *Int J Epidemiol*. 2021;49:2058-2064. doi:10.1093/ije/dyz132
19. Arnold KF, Davies V, de Kamps M, et al. Reflection on modern methods: generalized linear models for prognosis and intervention-theory, practice and implications for machine learning. *Int J Epidemiol*. 2021;49:2074-2082. doi:10.1093/ije/dyaa049
20. Xie Y, Brand JE, Jann B. Estimating heterogeneous treatment effects with observational data. *Sociol Methodol*. 2012;42:314-347. doi:10.1177/0081175012452652
21. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4:1. doi:10.1186/2046-4053-4-1
22. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169:467-473. doi:10.7326/M18-0850
23. Granholm A, Kaas-Hansen BS, Kjær M-BN, et al. Patient-important outcomes other than mortality in recent ICU trials: protocol for a scoping review. *Acta Anaesthesiol Scand*. 2021;65:1002-1007. doi:10.1111/aas.13937
24. Granholm A, Anthon CT, Kjær M-BN, et al. Patient-important outcomes other than mortality in contemporary ICU trials: a scoping review. *Crit Care Med*. 2022. doi:10.1097/CCM.0000000000005637
25. Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. & Welch, V., editors. 2022. Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022) (Cochrane, 2022).
26. Anthon CT, Granholm A, Perner A, Laake JH, Møller MH. No firm evidence that lack of blinding affects estimates of mortality in randomized clinical trials of intensive care interventions: a systematic review and meta-analysis. *J Clin Epidemiol*. 2018;100:71-81. doi:10.1016/j.jclinepi.2018.04.016
27. Anthon CT, Granholm A, Perner A, Laake JH, Møller MH. Overall bias and sample sizes were unchanged in ICU trials over time: a meta-epidemiological study. *J Clin Epidemiol*. 2019;113:189-199. doi:10.1016/j.jclinepi.2019.05.021
28. The Book of OHDSI. *Observational Health Data Sciences and Informatics [e-book]*. 1st ed. Independently published; 2021.
29. HL7. Index - FHIR v4.3.0. 2022. https://hl7.org/fhir//index.html
30. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi L, Mark R. 2020. MIMIC-IV (version 1.0). 10.13026/s6n6-xd98.
31. Pollard TJ, Johnson AEW, Raffa JD, et al. The eICU collaborative research database, a freely available multi-center database for critical care research. *Sci Data*. 2018;5:180178. doi:10.1038/sdata.2018.178
32. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37-46. doi:10.1177/001316446002000104
33. Rhodes W. Heterogeneous treatment effects: what does a regression estimate? *Eval Rev*. 2010;34:334-361. doi:10.1177/0193841X10372890

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kaas-Hansen BS, Granholm A, Anthon CT, et al. Causal inference for planning randomised critical care trials: Protocol for a scoping review. *Acta Anaesthesiol Scand*. 2022;66(10):1274-1278. doi:10.1111/aas.14142