

Common pitfalls in critical care research

Bruno Adler Maccagnan Pinheiro Besen^{1,2,3} , Antônio Paulo Nassar Jr^{1,3} , Juliana Carvalho Ferreira^{3,4} , Otavio Ranzani^{4,5} 

¹ Instituto D'Or de Pesquisa e Ensino - São Paulo (SP), Brazil.

² Postgraduate Program in Medical Sciences, Department of Internal Medicine, Faculdade de Medicina, Universidade de São Paulo - São Paulo (SP), Brazil.

³ Intensive Care Unit, A.C. Camargo Cancer Center - São Paulo (SP), Brazil.

⁴ Division of Pulmonology, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo - São Paulo (SP), Brazil.

⁵ Barcelona Institute for Global Health - Barcelona, Spain.

INTRODUCTION

Critical care research is characterized by many opportunities due to the data-rich environment of the intensive care unit (ICU): physiological monitoring derangements, patients' prior characteristics, acute organ dysfunction, metabolic derangements, and many laboratory and organ support measurements. Many ICUs integrate, in real-time, monitors, laboratory results and charts, allowing high time-resolution and completeness, providing large, open-sourced, databases.^(1,2) While this leads to a scientifically rich environment, it may cloud scientific thinking and lead to data-driven studies without clear hypotheses. Here, we will discuss common pitfalls researchers should be aware of when conducting critical care research.

Lack of a clear, relevant research question

A clear research question is a tenet for the planning, execution, and interpretation of the results, and it will pave the way to avoid many pitfalls. The underlying theory must guide the research, not the statistical results. Although it seems trivial, even experienced researchers may dive into a loophole due to an unclear research question. Research question clarity involves two main issues. First, the researcher must be sure of their primary aim: is it mainly descriptive, causal, or predictive? Second, the researcher should have a clear theoretical model of how the intervention works, in case of causal questions, or how a new diagnostic or predictive marker will fit in clinical practice, in case of diagnostic and prognostic research. Mixing concepts is very common and clouds scientific thinking.

Definition of inclusion and exclusion criteria

In clinical trials, much thought is put into inclusion and exclusion criteria to adequately sample the target population. This rationale must be followed for other study designs too and it is dependent on the study aim: for causal questions, exclusion criteria should mirror those of a pragmatic clinical trial. For prognosis and diagnosis questions, the study population should mirror the one that would have the test applied in practice. Authors should avoid excluding patients due to missing data, which should be appropriately dealt with.⁽³⁾

Over reliance on statistical significance instead of estimation

Traditional teaching in healthcare research emphasizes null hypothesis significance testing and p value interpretation. Researchers must focus, however, on estimation.⁽⁴⁾ Quantitative research is about estimating a given quantity of interest. It could be the prevalence of a condition for descriptive research, an estimate of risk ratios for a causal question, or the risk of an outcome for prognosis studies. This does not mean statistical significance is doomed, but that it must be presented with point estimates and confidence intervals and, sometimes, not even conducted depending on the scenario.⁽⁵⁾

Association of physiological derangements with death

Although critical care practice is intimately related to avoiding death, authors should avoid the common shortcut of finding an association between physiological derangements and mortality. Essentially, all physiological derangements will be to some extent associated with mortality (i.e., those who are most seriously ill are more likely to die), but these are commonly only prognostic associations. Well-thought research questions may need a demonstration of this association at the outstart of a line of research, but authors should consider potential causal paths amenable to treatment when moving forward.

Variable selection for multivariable models

Variable selection must be aligned with the research question. For causal questions, variable selection should be based on subject-matter knowledge by drawing conceptual models, such as causal directed acyclic graphs (DAGs).^(6,7) For predictive studies, as the aim is to estimate a probability, the causal relationship is not essential, but the outcome explainability.⁽⁸⁾ Nevertheless, for predictive questions, the authors must consider what variables are available at the point of care. These variables should not be a surrogate of what is meant to be predicted, which would lead to overly optimistic estimates of the model and poor applicability. Mixing concepts of prediction variable selection with DAG-based variable selection for causal questions may lead to low-value data presentation.

Table 2 Fallacy

The Table 2 Fallacy is an underrecognized issue in observational studies with multivariable models.⁽⁹⁾ Except in particular circumstances of prognostic research in which researchers aim to find an association of a combination of variables with clinical outcomes, not all coefficients of a multivariable model should be presented. In causal research, variables included in a multivariable model for adjustment should not have their results presented since *their* “effect estimate” is not estimated by *that* model. Each set of variables included in a model was so because they were necessary to address that research

question. Furthermore, even in prognostic research, the association of a new prognostic variable with outcome should be evaluated by accounting for known, observed prognostic variables in the model, regardless of hypothesis testing in that given dataset.

No pre-registration

Registration of clinical trials is a standard practice to avoid publication bias and undescribed modifications to study procedures during study roll-out. Systematic reviews must also have a registered protocol to allow *a-priori* hypothesis specification.⁽¹⁰⁾ Protocol deviations must be justified when needed. Absent pre-registration might preclude their publication. While observational studies still do not need pre-registration, authors should consider doing so or making the analysis protocol publicly available.

Abstracts without estimates

Abstracts are the first impression of a manuscript. In an abstract, a single main clear objective should be answered with the conclusion of the manuscript, but the focus should be on a brief description of methods and the main results’ presentation. The results need to be presented with estimates and their confidence intervals. Authors should avoid presenting only p values.

New methods are not always better methods

While as scientists we desire better methods to answer our research questions, new methods are not necessarily better methods. For example, the net reclassification improvement arose a decade ago as a new method for biomarker research, but it did not pass the test of time.⁽¹¹⁾ While Bayesian analysis,⁽¹²⁾ machine learning, and artificial intelligence^(13,14) are in the spotlight, they should not drive the research question.

CONCLUSION

From this viewpoint, we presented a few issues that we believe are important for critical care researchers. Figure 1 provides a contrast of good *versus* not-so-good practices in an observational study for further reference, which can be applied to a recently published systematic review.⁽¹⁵⁾

Concept	A good example	A not so good example
1. Clear research question 	Is awake prone positioning associated with reduced risk of intubation among COVID-19 patients with respiratory failure?	What are the risk factors associated with death and intubation in COVID-19? What is the role of awake prone positioning?
2. Well defined sample 	Patients admitted with COVID-19 who need oxygen supplementation and who do not need immediate intubation	All patients admitted to the hospital with COVID-19 (regardless of any criterion for respiratory failure)
3. Null-hypothesis significance testing 	Effect estimates and confidence intervals are presented followed by p values	p values are presented in tables and/or text without accompanying effect estimates
4. Physiological derangements and death 	No testing of variables other than the exposure with the outcome	Testing all variables in the database with mortality to create a table and interpret their results
5. Variable selection 	Variables selected based on confounding consideration through a DAG	Variable selection based on p-values < 0.20 from univariate analysis and stepwise selection procedures in multivariable models
6. Table 2 fallacy 	Effect estimates of variables included for adjustment are not presented. Table 2 presents results for sensitivity analyses or secondary outcomes instead of confounders variables effect estimates	Presentation of a table 2 which includes awake prone positioning effect estimate and of the variables included for adjustment, interpreting selected confounder variables with causal claims
7. Pre-registration 	Statistical analysis plan and code presented	Not considered
8. Abstracts need estimates 	Presentation with odds ratios or risk ratios, 95% confidence intervals and p values	Presentation of only p-values in the Abstract focusing on NHST
9. New methods are not always better methods 	Presentation with logistic regression. New methods used if applicable to improve causal inference claims	Use of a propensity score matching procedure without considering previous concepts

Figure 1 - Examples of a contrast of a good (and not so good) hypothetical observational study about the effects of awake prone positioning in hospitalized COVID-19 patients with respiratory failure.

DAG - directed acyclic graph; NHST - null hypothesis significance testing.

Publisher's note

Conflicts of interest: None.

Submitted on October 28, 2024

Accepted on November 8, 2024

Corresponding author:

Bruno Adler Maccagnan Pinheiro Besen
Instituto D'Or de Pesquisa e Ensino
Avenida Brigadeiro Luís Antônio, 5001, 2º andar - Jardim Paulista
Zip code: 01401-002 - São Paulo (SP), Brazil
E-mail: brunobesen@yahoo.com.br

Responsible editor: Thiago Costa Lisboa 

REFERENCES

- Johnson AE, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10(1):1.
- Pollard TJ, Johnson AE, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data*. 2018;5(1):180178.
- Vesin A, Azoulay E, Ruckly S, Vignoud L, Rusinovà K, Benoit D, et al. Reporting and handling missing values in clinical studies in intensive care units. *Intensive Care Med*. 2013;39(8):1396-404.
- Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of confounding and reporting of results in causal inference studies. Guidance for authors from editors of respiratory, sleep, and critical care journals. *Ann Am Thorac Soc*. 2019;16(1):22-8.
- Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*. 2016;31(4):337-50.
- Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. *J Clin Epidemiol*. 2022;142:264-7.
- Fumo-dos-Santos C, Ferreira JC. Dealing with confounding in observational studies. *J Bras Pneumol*. 2023;49(4):e20230281.
- Leisman DE, Harhay MO, Lederer DJ, Abramson M, Adjei AA, Bakker J, et al. Development and reporting of prediction models: guidance for authors from editors of respiratory, sleep, and critical care journals. *Crit Care Med*. 2020;48(5):623-33.
- Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *Am J Epidemiol*. 2013;177(4):292-8.
- Adigbli D, Yang L, Hammond N, Annane D, Arabi Y, Bilotta F, et al. Intensive glucose control in critically ill adults: a protocol for a systematic review and individual patient data meta-analysis. *Crit Care Sci*. 2023;35(4):345-54.

11. Pepe MS, Fan J, Feng Z, Gerds T, Hilden J. The Net Reclassification Index (NRI): a misleading measure of prediction improvement even with independent test data sets. *Stat Biosci.* 2015;7(2):282–95.
12. Zampieri FG, Damiani LP, Biondi RS, Freitas FG, Veiga VC, Figueiredo RC, et al.; BRICNet. Effects of balanced solution on short-term outcomes in traumatic brain injury patients: a secondary analysis of the BaSICS randomized trial. *Rev Bras Ter Intensiva.* 2022;34(4):410-7.
13. Ururahy RD, Gallo CA, Besen BA, Carvalho MT, Ribeiro JM, Zigaib R, et al. Bedside clinical data subphenotypes of critically ill COVID-19 patients: a cohort study. *Rev Bras Ter Intensiva.* 2021;33(2):196-205.
14. Santana A, Prestes GD, Silva MD, Girardi CS, Silva LD, Moreira JC, et al. Identification of distinct phenotypes and improving prognosis using metabolic biomarkers in COVID-19 patients. *Crit Care Sci.* 2024;36:e20240028en.
15. Vásquez-Tirado GA, Meregildo-Rodríguez ED, Asmat-Rubio MG, Salazar-Castillo MJ, Quispe-Castañeda CV, Cuadra-Campos MD. Conscious prone positioning in nonintubated COVID-19 patients with acute respiratory distress syndrome: systematic review and meta-analysis. *Crit Care Sci.* 2024;36:e20240176en.