LETTER TO THE EDITOR

DOI: https://doi.org/10.5114/ait/221640

How much is enough? Reconsidering the role of sample size in critical care research

Karol P. Steckiewicz

Department of Anesthesiology and Intensive Therapy, Faculty of Medicine, Medical University of Gdansk, Poland
Department of Anesthesiology and Intensive Therapy, University Clinical Centre, Gdansk, Poland

Anaesthesiol Intensive Ther 2026; 58: e82–e83

Received: 20.04.2026, **accepted:** 09.05.2026

Dear Editor,

The dismissal of small-sample intensive care unit (ICU) studies conflates sample size with methodological quality. As Ioannidis [1] famously argued, "most published research findings are false" in part due to biases that are not mitigated by increasing sample size alone. This concern is echoed in the broader methodological literature, where overreliance on *p*-values and power calculations has been widely criticized [2]. I argue that in critical care, sample size must be judged against the clinical, legal, and statistical context of the question asked. Sample size is a function of effect size, variance, and design assumptions, not an intrinsic marker of study quality.

Three constraints make large enrolment in the ICU structurally difficult, not merely inconvenient. First, the population of critically ill patients shows profound physiological heterogeneity, with trajectories that shift within hours. For rare syndromes or complex phenotypes, single-centre target populations may remain insufficient over years of recruitment, and multicentre designs are not always feasible. Second, the legal framework for consent-creates asymmetric research conditions across jurisdictions. Incapacity to consent (e.g. from sedation, delirium, encephalopathy, or mechanical ventilation) is the rule, not the exception. In the United States, 21 CFR 50.24 permits Exception From Informed Consent (EFIC) for emergency research under community consultation and IRB oversight. The EU Clinical Trials Regulation (No 536/2014) allows enrolment without prior consent in emergencies, but implementation across member states is heterogeneous and procedurally complex, producing delays in time-critical conditions such as septic shock or cardiac arrest even where deferred consent is legally available. These disparities have direct, quantifiable effects on enrolment. Third, the statistics. In the two-proportion framework, required *N* is inversely proportional to the square of the effect size. With a control event rate of 40%, an absolute risk reduction (ARR) of 15–20% reaches adequate power with roughly 50–120 patients per arm. A 20% ARR in 28-day mortality corresponds to a number needed to treat of 5, consistent with the magnitude of established ICU interventions, such as prone positioning in severe ARDS (ARR ≈17% in PROSEVA [3]). For continuous outcomes, required *N* depends on the ratio of variance to the minimum clinically important difference. In ICU populations, where physiological variability is typically high, this ratio inflates sample-size requirements even when the intervention effect is substantial.

Small studies have repeatedly produced signals that shaped definitive trials. Amato *et al*. [4] ($N = 53$) first demonstrated a mortality reduction with a lung-protective strategy in acute respiratory distress syndrome (ARDS), later confirmed by the ARDSNet ARMA trial [5] ($N = 861$), which established ventilation at 6 mL kg^{-1} predicted body weight as standard of care. The Rivers *et al*. [6] single-centre trial of early goal-directed therapy in septic shock

CORRESPONDING AUTHOR:

Karol P. Steckiewicz, Department of Anesthesiology and Intensive Therapy, Faculty of Medicine, Medical University of Gdansk, 17 Smoluchowskiego St., 80-214 Gdansk, Poland,
e-mail: karol.steckiewicz@gumed.edu.pl

(*N* = 263) transformed resuscitation practice for over a decade, though its mortality benefit was later relativized by ProCESS [7], ARISE [8], and ProMISe [9] showing that positive trials are hypothesis-generating, not verdict-delivering [10].

I propose that reviewers and editors replace numerical thresholds with a contextual appraisal:

1. Is the sample size justified by the primary outcome, design assumptions, and objectives?
2. Are absolute and relative effect estimates presented with confidence intervals?
3. Is the magnitude of the reported effect biologically and clinically plausible, and consistent with prior evidence?
4. Are regulatory and structural constraints on consent reported in enough detail to interpret recruitment?
5. Would a substantially larger trial have been feasible given the clinical, legal, and logistical setting?

The first three questions operationalise established reporting and evidence-grading principles (CONSORT 2025 [11]; GRADE [12]). The latter two, on jurisdictional consent reporting and feasibility, are not codified in current reporting standards but engage design challenges recognised by the critical care trialist community [13].

A study of 30–70 patients per arm can be adequate to detect the effect sizes that matter most in high-mortality ICU populations, provided the effect is real and the design is sound. The question, not the *N*, should be the primary criterion of scientific value.

ACKNOWLEDGEMENTS

1. Assistance with the article: None.
2. Financial support and sponsorship: None.
3. Conflicts of interest: None.
4. Presentation: None.
5. Description of the use of AI: LLM (Claude, Opus 4.6) was used for language polishing and formatting; all ideas, arguments, and cited evidence were developed independently by the author.

REFERENCES

1. Ioannidis JPA. Why most published research findings are false. PLoS Med 2005; 2: e124. DOI: 10.1371/journal.pmed.0020124.
2. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. Am Stat 2016; 70: 129-133. DOI: 10.1080/00031305.2016.1154108.
3. Guérin C, Reignier J, Richard JC, Beuret P, Gacouin A, Boulain T, et al. Prone positioning in severe acute respiratory distress syndrome. N Engl J Med 2013; 368: 2159-2168. DOI: 10.1056/NEJMoa1214103.
4. Amato MBP, Barbas CSV, Medeiros DM, Magaldi RB, Schettino GP, Lorenzi-Filho G, et al. Effect of a protective-ventilation strategy on mortality in the acute respiratory distress syndrome. N Engl J Med 1998; 338: 347-354. DOI: 10.1056/NEJM199802053380602.
5. Acute Respiratory Distress Syndrome Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. N Engl J Med 2000; 342: 1301-1308. DOI: 10.1056/NEJM200005043421801.
6. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med 2001; 345: 1368-1377. DOI: 10.1056/NEJMoa010307.
7. The ProCESS Investigators. A randomized trial of protocol-based care for early septic shock. N Engl J Med 2014; 370: 1683-1693. DOI: 10.1056/NEJMoa1401602.
8. The ARISE Investigators and the ANZICS Clinical Trials Group. Goal-directed resuscitation for patients with early septic shock. N Engl J Med 2014; 371: 1496-1506. DOI: 10.1056/NEJMoa1404380.
9. Mouncey PR, Osborn TM, Power GS, Harrison DA, Sadique MZ, Grieve RD, et al. Trial of early, goal-directed resuscitation for septic shock. N Engl J Med 2015; 372: 1301-1311. DOI: 10.1056/NEJMoa1500896.
10. The PRISM Investigators. Early, goal-directed therapy for septic shock – a patient-level meta-analysis. N Engl J Med 2017; 376: 2223-2234. DOI: 10.1056/NEJMoa1701380.
11. Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. BMJ 2025; 389. DOI: 10.1136/bmj-2024-081123.
12. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence – imprecision. J Clin Epidemiol 2011; 64: 1283-1293. DOI: 10.1016/j.jclinepi.2011.01.012.
13. Harhay MO, Casey JD, Clement M, Collins SP, Gayat É, Gong MN, et al. Contemporary strategies to improve clinical trial design for critical care research: insights from the First Critical Care Clinical Trialists Workshop. Intensive Care Med 2020; 46: 930-942. DOI: 10.1007/s00134-020-05934-6.