WHAT'S NEW IN INTENSIVE CARE

Should we rely on trials with disease- rather than patient-oriented endpoints?

Harm-Jan de Grooth[1,2]* , Jean-Jacques Parienti[3,4] and Heleen M. Oudemans-van Straaten[1]

Introduction

Should I change practice on the basis of this study [1]? Which primary endpoint should I choose to compute the sample size of my trial [2, 3]? Answering these questions requires a critical appraisal of study endpoints by both researchers and clinicians. In this report, we analyze the use of disease-oriented endpoints (such as organ dysfunction scores) in intensive care (IC) trials, and we briefly review the pitfalls of extrapolating disease-oriented endpoints to real patient benefit.

Disease-oriented endpoints in intensive care trials

The goal of IC research is to improve patients' health in ways that matter. Study endpoints that are important to patients, such as quality of life or survival, are referred to as "patient-oriented" endpoints because their relevance is self-evident and unambiguous [1, 2]. However, randomized trials often use "disease-oriented" endpoints, which correlate with patient-oriented benefit, but do not matter directly and unequivocally to patients [4].

Reasons to choose disease-oriented primary endpoints are that such endpoints generally require less patients and may be more sensitive indicators of treatment effects than survival. Examples of disease-oriented endpoints in IC research are organ failure scores, length of stay, time to shock reversal, or ventilator-free days. Although these endpoints have face validity, they leave room for ambiguity with respect to real patient benefit. A therapy may decrease organ failure without improving survival, suggesting that the new therapy worked, but did not help the patient survive.

We analyzed the primary endpoints of IC-related randomized trials published in five critical care journals and three high-impact medical journals over the past 15 years. The methods, baseline characteristics, and additional results are available as electronic supplementary material.

The use of disease-oriented primary endpoints significantly increased over time in trials with more than 200 patients and in trials published in high-impact journals (Fig. 1). For trials with more than 500 patients, disease-oriented endpoints were more prevalent than patient-oriented endpoints for the first time in 2016 (data not shown). Trials with disease-oriented primary endpoints more often reported positive results (i.e., *p* values less than 0.05) than trials with patient-oriented endpoints. Consequently, the question "Should we rely on trials with disease- rather than patient-oriented endpoints?" is critical.
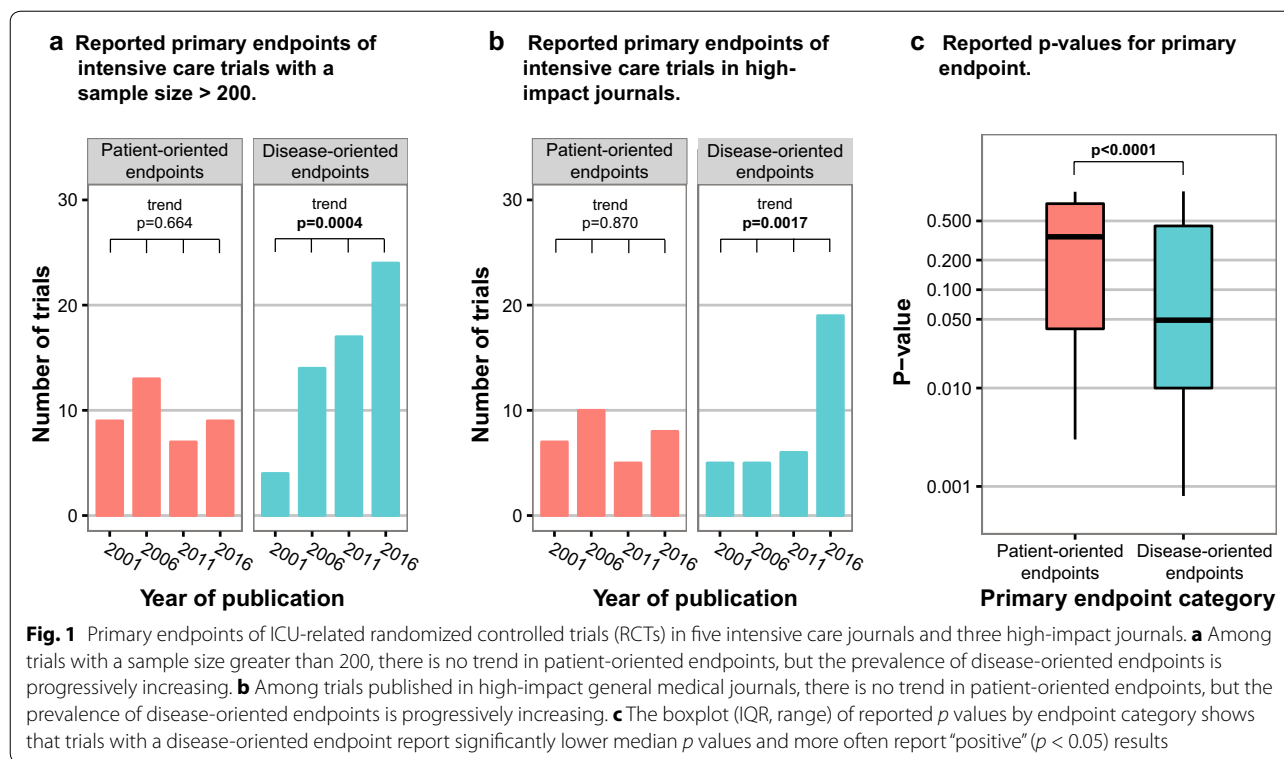
From disease-oriented endpoints to patient-oriented benefit

Several guidelines on the validation of disease-oriented endpoints have been published [1, 5]. Yet some recommendations are difficult to translate to IC research, which is characterized by complex syndromes (such as sepsis) rather than focused diseases (such as ischemic stroke). Consequently, disease-oriented endpoints are often broad and multifaceted (such as organ failure scores) rather than proximal surrogates for a specific clinical outcome (such as LDL-cholesterol as a surrogate for cardiovascular risk). No Food and Drug Administration approved surrogate endpoints are therefore currently used in IC research.

In general, a relevant endpoint must satisfy three criteria [5]: there must be biological plausibility that improvement in the disease-oriented endpoint will

*Correspondence: h.degrooth@vumc.nl
[1] Department of Intensive Care, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands
Full author information is available at the end of the article

Fig. 1 Primary endpoints of ICU-related randomized controlled trials (RCTs) in five intensive care journals and three high-impact journals. **a** Among trials with a sample size greater than 200, there is no trend in patient-oriented endpoints, but the prevalence of disease-oriented endpoints is progressively increasing. **b** Among trials published in high-impact general medical journals, there is no trend in patient-oriented endpoints, but the prevalence of disease-oriented endpoints is progressively increasing. **c** The boxplot (IQR, range) of reported *p* values by endpoint category shows that trials with a disease-oriented endpoint report significantly lower median *p* values and more often report "positive" ($p < 0.05$) results

cause improvement in true patient benefit; there must be a well-established observational association; and there must be evidence from intervention studies that the disease-oriented endpoint adequately captures the treatment effects on patient-oriented outcomes.

The capture criterion is especially important in IC research because the causal pathways between disease-oriented endpoints and real patient-oriented benefit are seldom clear and linear [6]. For example, the relationship between ventricular ectopy and sudden cardiac death after myocardial infarction is both plausible and statistically significant, but therapies that reduce ventricular ectopy after myocardial infarction (a disease-oriented endpoint) may paradoxically increase mortality [7]. For the same reason, a composite endpoint with disease- and patient-oriented components may be difficult to interpret when the treatment has different effects on the individual components.

Three complementary approaches improve our understanding of the relationship between disease- and patient-oriented endpoints in IC research.

Firstly, a critical appraisal of the literature may reveal vulnerabilities in the hypothesized causal chain between a disease-oriented endpoint and patient-oriented outcomes. For example, even though oxygenation and survival are both plausibly linked and statistically associated in acute respiratory distress syndrome (ARDS), there is no evidence that oxygenation impairment is the most

important driver of mortality [8]. It is therefore not surprising that, in ARDS trials, treatments that improve oxygenation (a frequently used endpoint) do not necessarily improve survival [9].

Secondly, a meta-analysis of the association between a disease-oriented endpoint and patient-oriented outcomes may elucidate the responsiveness and the reliability of the disease-oriented endpoint [10, 11]. We have recently shown in a meta-analysis of 87 RCTs that, on average, treatments that improved the Sequential Organ Failure Assessment (SOFA) score on a fixed day after randomization did not improve mortality. In contrast, treatments that improved delta-SOFA score (the trajectory from baseline) did improve mortality [12]. This shows that careful calibration is needed to make reliable inferences about patient-oriented benefit.

Thirdly, several statistical techniques can be used to analyze how well a disease-oriented endpoint captures the treatment effects on a patient-oriented outcome within a specific trial [13–15]. One such measure is the "proportion explained" (PE), which refers to the proportion of the treatment effect on a patient-oriented outcome explained by the treatment effect on the (primary) disease-oriented endpoint [14, 15]. A low PE may indicate that the disease-oriented endpoint fails to capture treatment effects on the patient-oriented endpoint (inappropriate disease-oriented endpoint) or that the patient-oriented endpoint is largely explained by treatment

allocation alone (irrelevant disease-oriented endpoint). Within-trial validation requires data which is generally only available to investigators. Therefore, an "endpoint validity statement" in a trial report (e.g., "the treatment effect on ventilator-free days explained 76% of the treatment effect on mortality") could improve the relevance of trials with disease-oriented primary endpoints.

In conclusion, disease-oriented primary endpoints have become progressively more prevalent in large and high-impact IC trials. However, ample evidence exists to make clinicians cautious before embracing any new therapy based on studies showing improvements in disease-oriented endpoints that have not been validated against patient-oriented outcomes. To evaluate the relevance of disease-oriented endpoints, we suggest to critically appraise the literature, to study the responsiveness between disease- and patient-oriented endpoints, and to report the PE in clinical trials. These strategies will help clinicians to be better informed about both the efficacy and the effectiveness of a treatment.

Electronic supplementary material

Author details

[1] Department of Intensive Care, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands. [2] Department of Anesthesiology, VU University Medical Center, Amsterdam, The Netherlands. [3] Unité de Biostatistique et de Recherche Clinique, Centre Hospitalier Universitaire de Caen, Avenue de la côte de Nacre, CS 30001, 14033 Caen, France. [4] EA4655 "Risques microbiens", Faculté de Médecine, Université de Caen Normandie, Caen, France.

Compliance with ethical standards

Conflicts of interest

Open Access

References

1. Bucher HC, Guyatt GH, Cook DJ, Evidence-Based Medicine Working Group et al (1999) Users' guides to the medical literature: XIX. Applying clinical trial results A. How to use an article measuring the effect of an intervention on surrogate end points. JAMA 282:771–778. doi:10.1001/jama.282.8.771
2. Tibby SM (2010) Does PELOD measure organ dysfunction…and is organ function a valid surrogate for death? Intensive Care Med 36:4–7. doi:10.1007/s00134-009-1712-0
3. Kellum JA, Zarbock A, Nadim MK (2017) What endpoints should be used for clinical studies in acute kidney injury? Intensive Care Med. doi:10.1007/s00134-017-4732-1
4. Fleming TR, DeMets DL (1996) Surrogate end points in clinical trials: are we being misled? Ann Intern Med 125:605–613
5. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (1998) ICH harmonized tripartite guideline: statistical principles for clinical trials E9
6. Rubenfeld GD (2003) Surrogate measures of patient-centered outcomes in critical care. In: Angus DC, Carlet J (eds) Surviving intensive care. Springer, Berlin, pp 169–180
7. Echt DS, Liebson PR, Mitchell LB et al (1991) Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The cardiac arrhythmia suppression trial. N Engl J Med 324:781–788. doi:10.1056/NEJM199103213241201
8. Stapleton RD, Wang BM, Hudson LD et al (2005) Causes and timing of death in patients with ARDS. Chest 128:525–532. doi:10.1378/chest.128.2.525
9. Afshari A, Brok J, Møller AM, Wettto J (2010) Inhaled nitric oxide for acute respiratory distress syndrome (ARDS) and acute lung injury in children and adults. Cochrane Database Syst Rev. doi:10.1002/14651858.CD002787.pub2
10. Buyse M, Molenberghs G, Burzykowski T et al (2000) The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics 1:49–67. doi:10.1093/biostatistics/1.1.49
11. Daniels MJ, Hughes MD (1997) Meta-analysis for the evaluation of potential surrogate markers. Stat Med 16:1965–1982. doi:10.1002/(SICI)1097-0258(19970915)16:17<1965:AID-SIM630>3.0.CO;2-M
12. de Grooth H-J, Geenen IL, Girbes AR et al (2017) SOFA and mortality endpoints in randomized controlled trials: a systematic review and meta-regression analysis. Crit Care 21:38. doi:10.1186/s13054-017-1609-1
13. Freedman LS, Graubard BI, Schatzkin A (1992) Statistical validation of intermediate endpoints for chronic diseases. Stat Med 11:167–178
14. Cleophas TJ, Zwinderman AH (2012) Validating surrogate endpoints of clinical trials. In: Statistics applied to clinical studies. Springer, Berlin, pp 569–578
15. Parast L, McDermott MM, Tian L (2016) Robust estimation of the proportion of treatment effect explained by surrogate marker information. Stat Med 35:1637–1653. doi:10.1002/sim.6820