Outcomes and Statistical Power in Adult Critical Care Randomized Trials

Michael O. Harhay1, Jason Wagner2, Sarah J. Ratcliffe1, Rachel S. Bronheim3, Anand Gopal1, Sydney Green1, Elizabeth Cooney1, Mark E. Mikkelsen2, Meeta Prasad Kerlin1,2, Dylan S. Small4, and Scott D. Halpern1,2,3
1Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, 2Division of Pulmonary, Allergy, and Critical Care Medicine, 3Department of Medical Ethics and Health Policy at the Perelman School of Medicine, and 4Department of Statistics at the Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania

Abstract

Rationale: Intensive care unit (ICU)-based randomized clinical trials (RCTs) among adult critically ill patients commonly fail to detect treatment benefits.

Objectives: Appraise the rates of success, outcomes used, statistical power, and design characteristics of published trials.

Methods: One hundred forty-six ICU-based RCTs of diagnostic, therapeutic, or process/systems interventions published from January 2007 to May 2013 in 16 high-impact general or critical care journals were studied.

Measurement and Main Results: Of 146 RCTs, 54 (37%) were positive (i.e., the a priori hypothesis was found to be statistically significant). The most common primary outcomes were mortality (n = 40 trials), infection-related outcomes (n = 33), and ventilation related outcomes (n = 30), with positive results found in 10, 58, and 43%, respectively. Statistical power was discussed in 135 RCTs (92%); 92 cited a rationale for their power parameters. Twenty trials failed to achieve at least 95% of their reported target sample size, including 11 that were stopped early due to insufficient accrual/logistical issues. Of 34 superiority RCTs comparing mortality between treatment arms, 13 (38%) accrued a sample size large enough to find an absolute mortality reduction of 10% or less. In 22 of these trials the observed control-arm mortality rate differed from the predicted rate by at least 7.5%.

Conclusions: ICU-based RCTs are commonly negative and powered to identify what appear to be unrealistic treatment effects, particularly when using mortality as the primary outcome. Additional concerns include a lack of standardized methods for assessing common outcomes, unclear justifications for statistical power calculations, insufficient patient accrual, and incorrect predictions of baseline event rates.