We declare no competing interests. Outside of the submitted work, AA has received grants and personal fees from AstraZeneca and Menarini, and personal fees from Chiesi, Teva, and Novartis; RF has received grants from GlaxoSmithKline, Menarini, and MSD; BC has received personal fees from AstraZeneca, GlaxoSmithKline, Boehringer Ingelheim, Novartis, Chiesi, and Menarini, and other from AstraZeneca; and RR-R has received personal fees from Pearl Therapeutics, Boehringer Ingelheim, and AstraZeneca.

1 Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. *Am J Respir Crit Care Med* 2017; **195:** 557–82.

2 Tan WC, Bourbeau J, Hernandez P, et al. Exacerbation-like respiratory symptoms in individuals without chronic obstructive pulmonary disease: results from a population-based study. *Thorax* 2014; **69:** 709–17.

3 Beghé B, Verduri A, Roca M, Fabbri LM. Exacerbation of respiratory symptoms in COPD patients may not be exacerbations of COPD. *Eur Respir J* 2013; **41:** 993–95.

4 Scioscia G, Blanco I, Arismendi E, et al. Different dyspnoea perception in COPD patients with frequent and infrequent exacerbations. *Thorax* 2017; **72:** 117–21.

5 Bafadhel M, McKenna S, Terry S, et al. Acute exacerbations of chronic obstructive pulmonary disease: identification of biologic clusters and their biomarkers. *Am J Respir Crit Care Med* 2011; **184:** 662–71.

6 Jameson JL, Longo DL. Precision medicine—personalized, problematic, and promising. *N Engl J Med* 2015; **372:** 2229–34.

7 Noell G, Cosío BG, Faner R, et al. Multi-level differential network analysis of COPD exacerbations. *Eur Respir J* 2017; **50:** 1700075.

8 Pinto-Plata VM, Livnat G, Girish M, et al. Systemic cytokines, clinical and physiological changes in patients hospitalized for exacerbation of COPD. *Chest* 2007; **131:** 37–43.

9 Hurst JR, Donaldson GC, Perera WR, et al. Use of plasma biomarkers at exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006; **174:** 867–74.

10 Bozinovski S, Hutchinson A, Thompson M, et al. Serum amyloid A is a biomarker of acute exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2008; **177:** 269–78.

11 Thygesen K, Alpert JS, Jaffe AS, et al. Third universal definition of myocardial infarction. *Eur Heart J* 2012; **33:** 2551–67.

12 Shafiek H, Fiorentino F, Merino JL, et al. Using the electronic nose to identify airway infection during COPD exacerbations. *PLoS One* 2015; **10:** e0135199.

13 Anthonisen NR, Manfreda J, Warren CP, Hershfield ES, Harding GK, Nelson NA. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med* 1987; **106:** 196–204.

14 Simpson JL, Baines KJ, Horvat JC, et al. COPD is characterized by increased detection of Haemophilus influenzae, Streptococcus pneumoniae and a deficiency of Bacillus species. *Respirology* 2016; **21:** 697–704.

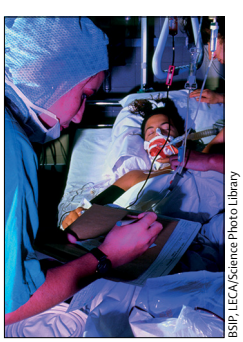
15 Bafadhel M, McKenna S, Terry S, et al. Blood eosinophils to direct corticosteroid treatment of exacerbations of chronic obstructive pulmonary disease: a randomized placebo-controlled trial. *Am J Respir Crit Care Med* 2012; **186:** 48–55.

Negative trials in critical care: why most research is probably wrong

Randomised controlled trials (RCTs) reduce the potential for allocation bias, and the results are considered by many to be immediately implementable in practice and form the backbone of clinical guidelines. However, a major limitation of RCTs in critical care is that the overwhelming majority of results are negative, meaning the intervention proved no better (or was worse) than the control. In 2008, it was reported that 62 (83%) of 75 RCTs investigating mortality in critical care showed no benefit from the intervention.[1] Of the positive RCTs, three of the results are now generally accepted, yielding a total of less than 5% reliably positive results. Similar results were reported in 2014 in the context of acute respiratory distress syndrome (ARDS).[2] At the time of writing, the proportion of positive results from the 20 most recent large pharmacological studies of both sepsis and ARDS is 5% (appendix).

On a statistical basis, it is unlikely that most negative RCTs represent fair conclusions. For example, if each hypothesis tested had the same a priori probability that was as low as a coin toss, we would expect 50% of RCTs to be positive. The likelihood that 19 out of 20 consecutive such coin tosses would be negative is less than 1 in half a million ($1 \div 2^{19}$). In fact, most a priori hypotheses should have substantially better than even odds of success, as it seems implausible that a hypothesis refined through extensive testing should have only the same chance of being effective as a coin toss. Bayesian probability reflecting the track record of positive RCTs in critical care, suggests a less than 1-in-20 chance that the next study will be positive (assuming only a 50:50 forecast).

The usual explanation for false-negative results is that patient numbers are too small (and the acceptable p-value is not small enough). But if the intervention is not matched to the patients being studied (eg, the biological target is absent) then the limitation is biological—not statistical—and insistence on greater numbers (or more stringent p-values) will have no effect. This is a major problem in RCTs in critical care, because the entry criteria into almost all studies of sepsis or ARDS (prototypic syndromes in critically ill patients) rely not on a specific diagnosis, but on consensus definitions. Although these definitions are excellent for screening (ie, have high sensitivity), they perform poorly as diagnostic tests (ie, have low specificity), which is unsurprising because sepsis and ARDS are biologically heterogenous syndromes: they are not diseases per se with singular mechanisms that are plausibly amenable to singular interventions. Failure



BSIP, LECA/Science Photo Library

Published Online
July 27, 2018
http://dx.doi.org/10.1016/S2213-2600(18)30279-0

See Online for appendix

to match the proposed intervention with a potentially responsive patient dilutes any therapeutic effect; and, if the intervention cannot plausibly help a given patient, it can only be neutral or harmful. Finally, designating some RCTs as pragmatic might sometimes imply that the underlying biological mechanism of action is not central.

There are important reasons to detect and prevent false-negative RCTs in critical care. The result of an RCT is seen as the definitive statement about whether an intervention works in patients; therefore, a false-negative RCT can doom a potentially useful therapy. Given the arguments above, it seems likely that potentially useful therapies have been discarded.

A preponderance of negative studies can have additional adverse impacts. Although negative studies can outline what does not work (or is not an improvement), progress in a field depends on positive studies because only by a sequence of positive studies can new mechanisms be elucidated. It is unlikely that the brightest minds will be attracted to a specialty that produces almost entirely negative trials. Furthermore, funding agencies might become concerned at the low success rate of critical care trials; and, the repetitive pattern of negative RCTs raises important issues regarding utility and ethics.

Several important steps could provide a way forward and out of the cycle of negative RCTs in critically ill patients. For example, although conventional consensus criteria are useful for preliminary screening, they might be insufficient to determine trial eligibility, in the same way that screening criteria would never be an adequate basis to study (or prescribe) potent therapies for breast or prostate cancer. A parallel can be drawn with recent advances in oncology catalysed by systematic approaches to identifying molecular tumour characteristics.[3] This advance upon clinical or histological characterisation has enabled targeted molecular therapies, some of which have radically improved patients' prospects. Similar approaches concentrating on disease mechanisms in subgroups of patients with ARDS and sepsis are beginning to yield important insights.[4]

Enrichment of studies in this way should not simply represent selection of sicker patients, but rather selection of patients with a strong likelihood of a response to the intervention (while patients who are unlikely to respond should be selected out). This approach could reduce study noise, sample size, and study-associated harm, and has already been useful in ARDS: an important (positive) study[5] of prone positioning randomised only patients who showed an initial positive response to prone positioning. In addition, latent class analysis has identified a third of ARDS patients with a hyperinflammatory phenotype, and reanalysis of a large negative RCT of simvastatin in ARDS using this approach suggested a benefit in the hyperinflammatory group.[6]

In summary, excessive numbers of negative (many false-negative) trials constitute a major problem in critical care medicine. Researchers should focus on identifying biological mechanisms in sepsis or ARDS for which testable therapies can be proposed. Designers of RCTs and clinical trials groups should ensure that proposed RCTs in critical care identify subgroups of patients that match the specific intervention being tested. RCTs that advance our understanding of mechanisms are essential. Only by taking this approach can we end the cycle of negative RCTs, and move forward to developing improved therapies for patients with life-threatening critical illnesses.

*John G Laffey, *Brian P Kavanagh*
Regenerative Medicine Institute (REMEDI) at the Centre for Research in Medical Devices (CÚRAM), Biomedical Sciences, National University of Ireland Galway, Galway, Ireland (JGL); School of Medicine, College of Medicine, Nursing, and Health Sciences, National University of Ireland Galway, Galway, Ireland (JGL); Department of Anaesthesia, Department of Physiology, and Interdepartmental Division of Critical Care, University of Toronto (JGL, BPK); Department of Critical Care Medicine, Hospital for Sick Children, Toronto, ON M5G 1X8, Canada (BPK)
brian.kavanagh@utoronto.ca

JGL is supported by grant funding from the Canadian Institutes of Health Research and from Science Foundation Ireland (16/FRL/3845). BPK is supported by grant funding from the Canadian Institutes of Health Research and holds the Dr Geoffrey Barker Chair in Critical Care Research.

1 Ospina-Tascón GA, Büchele GL, Vincent JL. Multicenter, randomized, controlled trials evaluating mortality in intensive care: doomed to fail? *Crit Care Med* 2008; **36:** 1311–22.
2 Tonelli AR, Zein J, Adams J, Ioannidis JP. Effects of interventions on survival in acute respiratory distress syndrome: an umbrella review of 159 published randomized trials and 29 meta-analyses. *Intensive Care Med* 2014; **40:** 769–87.
3 Drilon A, Laetsch TW, Kummar S, et al. Efficacy of larotrectinib in TRK fusion-positive cancers in adults and children. *N Engl J Med* 2018; **378:** 731–39.
4 Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *Lancet Respir Med* 2014; **2:** 611–20.
5 Guérin C, Reignier J, Richard JC, et al. Prone positioning in severe acute respiratory distress syndrome. *N Engl J Med* 2013; **368:** 2159–68.
6 Calfee CS, Delucchi KR, Matthay MA,et al. Consistent ARDS endotypes are identified using minimal data from a United Kingdom clinical trial. *Am J Respir Crit Care Med* 2017; **195:** A6749.