EDITORIAL

Is research from databases reliable? No

CrossMark

Anders Perner[1*] , Rinaldo Bellomo[2,3] and Morten Hylander Møller[1]

Introduction

The findings of clinical research inform clinicians, guideline committee members and policy-makers on how to provide the best care of patients at the lowest possible cost. It is of utmost importance that such research findings are trustworthy. If not, we may harm patients and waste resources by the use of interventions with an unfavourable balance between risk and benefit. The risk of harm is real as indicated in a recent systematic review of all interventions shown to affect mortality in critical care trials [1]. Among the 15 interventions that affected mortality, eight increased mortality. Many of these harmful interventions were in clinical use at the time of testing, including tight glycaemic control, hydroxyethyl starch and high-frequency oscillatory ventilation [2–4]. Faced with the need to identify and promote interventions that may help patients and with the similar need to identify and avoid interventions that may harm patients, intensivists have used a variety of approaches. One such approach has been based on the use of databases or registries. Such an approach, however, carries important risks. In particular, the reliability of findings from such investigations is open to challenge.

Why are estimates of intervention effects unreliable from databases and registries?

Empiric evidence suggests that the results of observational studies based on databases or registries are unreliable [5, 6]. Overall, the observational data may overestimate the benefits of interventions as compared to results of randomised clinical trials (RCTs) [5, 6], but harm may also be overestimated as observed in analyses of observational studies vs. placebo-controlled RCTs on digoxin in patients with heart failure [5]. The reasons for the biased results are multiple and often include the design of the databases/registries, the quality of the data, confounding by indication, and the design of research studies based on such databases (Table 1). Moreover, if the database being assessed is local (single centre), as is the case for small studies, it may be more detailed, but it will also lack external validity. On the other hand, if a large dataset is being used, data "missingness" will be common and quality controls will often be limited.

Can the problems be fixed?

No, probably not. In the digoxin case, the overestimation of harm was largest in unadjusted analyses (76% relative risk increase in mortality with the use of digoxin) [5], but in adjusted analyses including propensity scoring, the overestimation remained. Thus, the propensity matched analysis overestimated the relative risk of harm from digoxin by 18% [5]. Moreover, inferences about critical care interventions from studies of databases and registries may be particularly prone to bias because of multiple, time-dependent exposures and many confounding co-interventions in critical care. Competing risk is another major challenge as many critical ill patients experience adverse events, which may mimic harmful effects of interventions, e.g., bleeding and kidney and hepatic injury.

In addition to the above problems, database-driven research is often asked to do more than it can do. Like all observational research, however, it can never deliver inferences of causality. Unfortunately, associations are often presented with a language that suggests such causality. For example, expressions like biomarker or intervention X "has an impact on" or "has an independent effect on" or "modifies" or "changes" a given outcome are often used. Such expressions subtly suggest causality where none can be inferred to exist. Moreover, phrases like "after adjustment for confounders" or "independent

*Correspondence: anders.perner@regionh.dk
[1] Department of Intensive Care, Rigshospitalet,
University of Copenhagen, Copenhagen, Denmark
Full author information is available at the end of the article

Springer

Table 1 Problems with observational studies from databases or registries vs. randomised trials in critical care

	Database/registry studies	Randomised trials
The design of the database	Often generic or build for a purpose other than research (e.g., administration, billing or quality control)	Customised to answer the specific research question
The quality of the data	Often routinely collected data entered by clinicians or administrative staff with high risk of erroneous data entry and missing data	Often entered from source data by dedicated research staff following strict definitions and subjected to quality control including external monitoring
Confounding by indication and residual confounding	The inherent problem of observational studies of interventions—causality cannot be shown.	Randomisation controls for measured and unmeasured confounders
The design of the research studies	Often less detailed protocols, which rarely are registered or published	Detailed trial protocol, which are submitted to the authorities, registered and often published
The analyses	Often not predefined but data-driven without validation in an independent dataset	Often predefined according to a statistical analysis plan
The interpretation	Often oversold by the use of wording like "independent effect on" or "modifies" or "changes"	By construct the interpretation is more straight forward

association" or "propensity score adjusted impact" are often used. However, associations can only be adjusted for given available variables; many variables that influence outcome (patient frailty, doctor's judgement, the exact situation at the time of the intervention etc.) are actually not available. Thus, all adjustments are flawed by the very nature of database characteristics. Such associations as may be found, therefore, are only hypothesis generating and should never lead to practice change. It would be good for critical care clinicians to reflect that, so far, no studies of associations have led to the identification of an intervention, which was then subsequently confirmed to affect outcome in a large multicentre RCT.

Have databases and registries no place in critical care research?

High quality observational data are the main source for information about incidences, complications, outcomes, clinical practice and resource use and for the development of diagnostic criteria and prognostic scoring systems. Longitudinal databases can identify secular trends in the incidence and outcome of conditions of interest such as sepsis, trauma, acute kidney injury, and ARDS. There are many good and recent examples in critical care [7–10], which provide useful information for policy-making and resource allocations. Exploratory analyses of the associations of specific interventions and outcome and estimates regarding recruitment rates and trial size have value for the planning of RCTs [11]. Finally, databases can be used to identify illness severity adjusted potential performance outliers that may require investigation [12].

However, research from databases needs to evolve in term of analysis protocols. In particular, the pre-publication of statistical analysis plans is necessary to give the reader confidence that the analysis was truly driven by the desire to test a hypothesis, rather than by a data-mining exercise, with all the inherent biases of such an approach. Modern RCTs now demand such pre-publication of protocols [13, 14] and statistical analysis plans [15, 16]. Similar demands for observational database-driven investigations should be made by journals to help attenuate the flaws of database research [17].

Perspectives

Critical care has evolved tremendously over the last half-century. Many of the interventions that have formed and form critical care were introduced into clinical practice without evidence from RCTs and systematic reviews with overall low risk of bias. The lesson learned in recent years is that the use of interventions based on the results of biased research, including those of observational studies of databases or registries may at best be a waste of time and money, but we may also harm patients [1]. Thus, it

is imperative for the critical care community to conduct low risk of bias RCTs to answer clinical questions on the balance between the benefits and harms of common interventions to provide the best and most cost-effective care for critically ill patients.

Author details
[1] Department of Intensive Care, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark. [2] Department of Intensive Care, Austin Hospital, School of Medicine, The University of Melbourne, Melbourne, VIC, Australia. [3] Australian and New Zealand Intensive Care Research Centre, School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia.

References
1. Landoni G, Comis M, Conte M, Finco G, Mucchetti M, Paternoster G, Pisano A, Ruggeri L, Alvaro G, Angelone M, Bergonzi PC, Bocchino S, Borghi G, Bove T, Buscaglia G, Cabrini L, Callegher L, Caramelli F, Colombo S, Corno L, Del SP, Feltracco P, Forti A, Ganzaroli M, Greco M, Guarracino F, Lembo R, Lobreglio R, Meroni R, Monaco F, Musu M, Pala G, Pasin L, Pieri M, Pisarra S, Ponticelli G, Roasio A, Santini F, Silvetti S, Szekely A, Zambon M, Zucchetti MC, Zangrillo A, Bellomo R (2015) Mortality in multicenter critical care trials: an analysis of interventions with a significant effect. Crit Care Med 43:1559–1568
2. Finfer S, Chittock DR, Su SY, Blair D, Foster D, Dhingra V, Bellomo R, Cook D, Dodek P, Henderson WR, Hebert PC, Heritier S, Heyland DK, McArthur C, McDonald E, Mitchell I, Myburgh JA, Norton R, Potter J, Robinson BG, Ronco JJ (2009) Intensive versus conventional glucose control in critically ill patients. New Engl J Med 360:1283–1297
3. Perner A, Haase N, Guttormsen AB, Tenhunen J, Klemenzson G, Aneman A, Madsen KR, Moller MH, Elkjaer JM, Poulsen LM, Bendtsen A, Winding R, Steensen M, Berezowicz P, Soe-Jensen P, Bestle M, Strand K, Wiis J, White JO, Thornberg KJ, Quist L, Nielsen J, Andersen LH, Holst LB, Thormar K, Kjaeldgaard AL, Fabritius ML, Mondrup F, Pott FC, Moller TP, Winkel P, Wetterslev J (2012) Hydroxyethyl starch 130/0.42 versus Ringer's acetate in severe sepsis. New Engl J Med 367:124–134
4. Ferguson ND, Cook DJ, Guyatt GH, Mehta S, Hand L, Austin P, Zhou Q, Matte A, Walter SD, Lamontagne F, Granton JT, Arabi YM, Arroliga AC, Stewart TE, Slutsky AS, Meade MO (2013) High-frequency oscillation in early acute respiratory distress syndrome. New Engl J Med 368:795–805
5. Ziff OJ, Lane DA, Samra M, Griffith M, Kirchhof P, Lip GY, Steeds RP, Townend J, Kotecha D (2015) Safety and efficacy of digoxin: systematic review and meta-analysis of observational and controlled trial data. BMJ 351:h4451
6. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JP (2016) Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. BMJ 352:i493
7. Kaukonen KM, Bailey M, Suzuki S, Pilcher D, Bellomo R (2014) Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000–2012. JAMA 311:1308–1316
8. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM, Hotchkiss RS, Levy MM, Marshall JC, Martin GS, Opal SM, Rubenfeld GD, van der Poll T, Vincent JL, Angus DC (2016) The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA 315:801–810
9. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, Gattinoni L, van HF, Larsson A, McAuley DF, Ranieri M, Rubenfeld G, Thompson BT, Wrigge H, Slutsky AS, Pesenti A (2016) Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. JAMA 315:788–800
10. Azoulay E, Pickkers P, Soares M, Perner A, Rello J, Bauer PR, van de Louw A, Hemelaar P, Lemiale V, Taccone FS, Martin LI, Meyhoff TS, Salluh J, Schellongowski P, Rusinova K, Terzi N, Mehta S, Antonelli M, Kouatchet A, Barratt-Due A, Valkonen M, Landburg PP, Bruneel F, Bukan RB, Pene F, Metaxa V, Moreau AS, Souppart V, Burghi G, Girault C, Silva UVA, Montini L, Barbier F, Nielsen LB, Gaborit B, Mokart D, Chevret S (2017) Acute hypoxemic respiratory failure in immunocompromised patients: the Efraim multinational prospective cohort study. Intensive Care Med 43:1808–1819
11. Krag M, Perner A, Wetterslev J, Wise MP, Borthwick M, Bendel S, McArthur C, Cook D, Nielsen N, Pelosi P, Keus F, Guttormsen AB, Moller AD, Moller MH (2015) Prevalence and outcome of gastrointestinal bleeding and use of acid suppressants in acutely ill adult intensive care patients. Intensive Care Med 41:833–845
12. McClean K, Mullany D, Huckson S, van LA, Chavan S, Hicks P, Hart G, Paul E, Pilcher D (2017) Identification and assessment of potentially high-mortality intensive care units using the ANZICS Centre for Outcome and Resource Evaluation clinical registry. Crit Care Resusc 19:230–238
13. Krag M, Perner A, Wetterslev J, Wise MP, Borthwick M, Bendel S, Pelosi P, Keus F, Guttormsen AB, Schefold JC, Moller MH (2016) Stress ulcer prophylaxis with a proton pump inhibitor versus placebo in critically ill patients (SUP-ICU trial): study protocol for a randomised controlled trial. Trials 17:205
14. Hammond NE, Bellomo R, Gallagher M, Gattas D, Glass P, Mackle D, Micallef S, Myburgh J, Saxena M, Taylor C, Young P, Finfer S (2017) The Plasma-Lyte 148 v Saline (PLUS) study protocol: a multicentre, randomised controlled trial of the effect of intensive care fluid therapy on mortality. Crit Care Resusc 19:239–246
15. Billot L, Venkatesh B, Myburgh J, Finfer S, Cohen J, Webb S, McArthur C, Joyce C, Bellomo R, Rhodes A, Perner A, Arabi Y, Rajbhandari D, Glass P, Thompson K, Correa M, Harward M (2017) Statistical analysis plan for the Adjunctive Corticosteroid treatment in critically Ill patients with septic shock (ADRENAL) trial. Crit Care Resusc 19:183–191
16. Krag M, Perner A, Wetterslev J, Lange T, Wise MP, Borthwick M, Bendel S, Pelosi P, Keus F, Guttormsen AB, Schefold JC, Meyhoff TS, Marker S, Moller MH (2017) Stress ulcer prophylaxis in the intensive care unit trial: detailed statistical analysis plan. Acta Anaesthesiol Scand 61:859–868
17. Granholm A, Perner A, Krag M, Hjortrup PB, Haase N, Holst LB, Marker S, Collet MO, Jensen AK, Moller MH (2017) Simplified Mortality Score for the Intensive Care Unit (SMS-ICU): protocol for the development and validation of a bedside clinical prediction rule. BMJ Open 7:e015339