Intensive Care Med (2019) 45:118–121
https://doi.org/10.1007/s00134-018-5436-x

EDITORIAL

Is research from databases reliable? Yes

Jean-Francois Timsit[1,2*], Jerome Aboab[3] and Jean-Jacques Parienti[4,5]

Whatever the tools used, any reflection on clinical research must bring us to humility. Indeed, the study results often vary from one study to another, providing contradictory evidence that does little to answer our questions [1, 2]. As for the available tools and methodologies go, they are numerous, each with its own strengths and weaknesses.

Large multicenter randomized controlled trials (RCTs) are considered the "gold standard" for comparing two treatment strategies. Indeed, the randomization is the only way to balance at random known and unknown or unmeasured confounding factors. The causality in RCTs implies that the initial conditions of randomization remain respected until the assessment of the primary endpoint. Dropout and lack of double-blind interventions, frequently observed in RCTs, are conditions of loss of the exchangeability assumption of RCTs.

But more and more voices are being raised against the hegemony of RCTs [3]. Indeed, not all types of exposure can be randomized. In addition, RCTs are conducted in selected groups which may not be representative of the target population. Furthermore, RCTs only compare precisely defined interventions or exposures, and the results can not be extrapolated to other interventions. This is of importance when the control group is not a faithful mirror of the usual care.

Poor compliance to the assigned treatment is an obvious limitation in determining causal effects in experimental designs. While an "as treated" analysis can be useful, intention-to-treat analysis is preferred but can underestimate causal treatment effect [4]. Replication of the results by an independent randomized trial is mandatory to change practices; however, it is not always available, or feasible. In the setting of critical care medicine, "gold" is expensive and scarce [5]. Can we trust well-conducted and appropriately analyzed high-quality databases or are they simply compilations of clinical observations? We argue for the important role that non-randomized studies can play in clinical research.

Added value of research on clinical databases

Research on clinical databases should not replace RCTs. However, we think it may be instrumental in increasing the level of evidence or raising new hypotheses [6]. This opinion is comforted by the consistency between appropriately analyzed high-quality databases and randomized trials [7].

First, studies on databases could generate hypotheses about the impact of exposure or of the intervention. Second, by raising the level of uncertainty among clinicians, they can increase the likelihood of clinicians participating in a future RCT. Third, they provide a permanent infrastructure for designing large multicenter trials. Fourth, they can be used to confirm on a general unselected population of patients the generalizability of the results of RCTs which are most often performed in highly selected specialized teaching intensive care units (ICUs). Fifth, by including a larger sample size than RCTs, clinical databases provide an increased statistical power for evaluating patient-oriented outcomes such as mortality, while RCTs often evaluate disease-oriented outcomes [8].

Good-quality data are available and usable: from big data to high-quality research databases and prospective cohorts

The high-tech ICU generates a considerable amount of data from different data sources on a patient or departmental level [9]. However, these data are captured for many other reasons than research, mainly care, administrative, and cost purposes. The use of administrative databases for research may lead to bias and spurious

*Correspondence: jean-francois.timsit@aphp.fr
[1] Medicine and Infectious Diseases ICU (MI2), APHP Bichat Hospital, Paris, France
Full author information is available at the end of the article

For contrasting viewpoints, please go to https://doi.org/10.1007/s00134-018-5073-4 and https://doi.org/10.1007/s00134-018-5498-9.

Springer

conclusions. Second, automatic captures of signals from machines may not clearly differentiate artifacts from reliable pieces of information. However, the use of this kind of data is cheap and may be interesting in quickly generating some hypotheses. However, overconfidence in the data reliability, and an incomplete understanding of the underlying relationships, could lead to dubious results and spurious interpretations. Unmeasured confounders, missing data, and inherent noise can generate false positive associations [10].

In contrast, much data (case-mix, severity, clinical and biological variables, procedure used, etc.) are collected in routine practice and are used in almost all research projects. In intensive care, high-quality research databases already exist and have fueled the discussion in many areas such as vascular filling [11], sepsis [12], ARDS, or renal failure [13]. As an example, the OUTCOMEREA network has been using this principle for 20 years to create a high-quality database dedicated only to research purposes and which prospectively captures more than 200 variables each ICU day. Such a database has already been used to answer more than 100 research questions.

Thus the challenge could be to position the decision boundary between a system database adapted for the follow-up and treatment of all patients and a project database dedicated to answering a single clinical research question (Fig. 1) [2]. Projects such as MIMIC II (https://physionet.org/mimic2/, accessed 22 October 2018) in the Beth Israel Deaconess Medical Center in Boston or the EDS built by the Assistance Publique-Hôpitaux de Paris (https://recherche.aphp.fr/eds/, accessed 14 October 2018) are to create data marts for research purposes extracted from a data warehouse that contains all the clinical and biological electronic records, examinations, and prescriptions of a given set of patients in numerical, text, or image formats (compiled from all the hospital stays from 40 hospitals in the Parisian area). These projects are ongoing and require precise descriptions and definitions of each variable collected, and careful audits of the data captured before merging and analyses.

Causal inference: how could it be applied to non-randomized studies?

The intrinsic feature of randomized or experimental study design is to provide two or more groups which share the same risk of endpoint before treatment. Therefore, only the allocated intervention differs between groups and thus will influence the outcome, allowing one to draw a causal inference. This may not be the case in non-randomized studies, where the intervention is provided by choice rather than by chance, raising the flag of confounding factors. Confounding is a bias in the estimated measure of association that occurs when the primary exposure of interest is mixed up with some other factors associated with the outcome. In order for confounding to occur, the extraneous factor must be associated with both the primary exposure of interest and the outcome of interest. If the distribution of the extraneous factor is similar in the exposure groups being compared, then this factor will not cause confounding.

Two contemporary statistical methods can be used to approach causal inference: multivariable and propensity score analyses. Both generally yield similar estimates [14], although the latter confers several advantages. In particular, propensity score methods [15] offer balanced case-mix between groups when the propensity score is used as a matching or weighting variable through the use of standardized difference.

Comparing groups with similar baseline characteristics is necessary but not always sufficient for causal inference. While double-blind randomized design maintains the effect of randomization during follow-up, observational studies are open-label. Therefore, immortal time bias is another very common source of distortion in observational studies [16]. The use of time-dependent exposure in longitudinal cohorts can overcome this problem in the framework of marginal structural models [17]. Of course, all these advanced statistics rely on two strong hypotheses: that unmeasured confounding data and model misspecifications are absent. This means that high-quality database and close collaboration between intensivists and biostatisticians are required [18].

Conclusion: what should be the position of research using databases in ICU clinical research?

The exploitation of large structured data warehouses of medical patients allows us to identify certain therapeutic strategies and to raise new questions. In the field of pharmacovigilance, registries have been used to identify unexpected side effects which led to changes in the indications or conditions of use of drugs. More recently, in the field of imaging, pneumonia detection algorithms are reportedly challenging the performance of radiologists in the detection of diseases.

High-quality research databases are important tools that may find not only associations but also causal relationships in ICUs. If properly built and carefully maintained, such databases could be used alone in rare diseases, before RCTs in order to generate quantitative hypotheses, and after RCTs to confirm hypotheses and ensure their external generalizability.

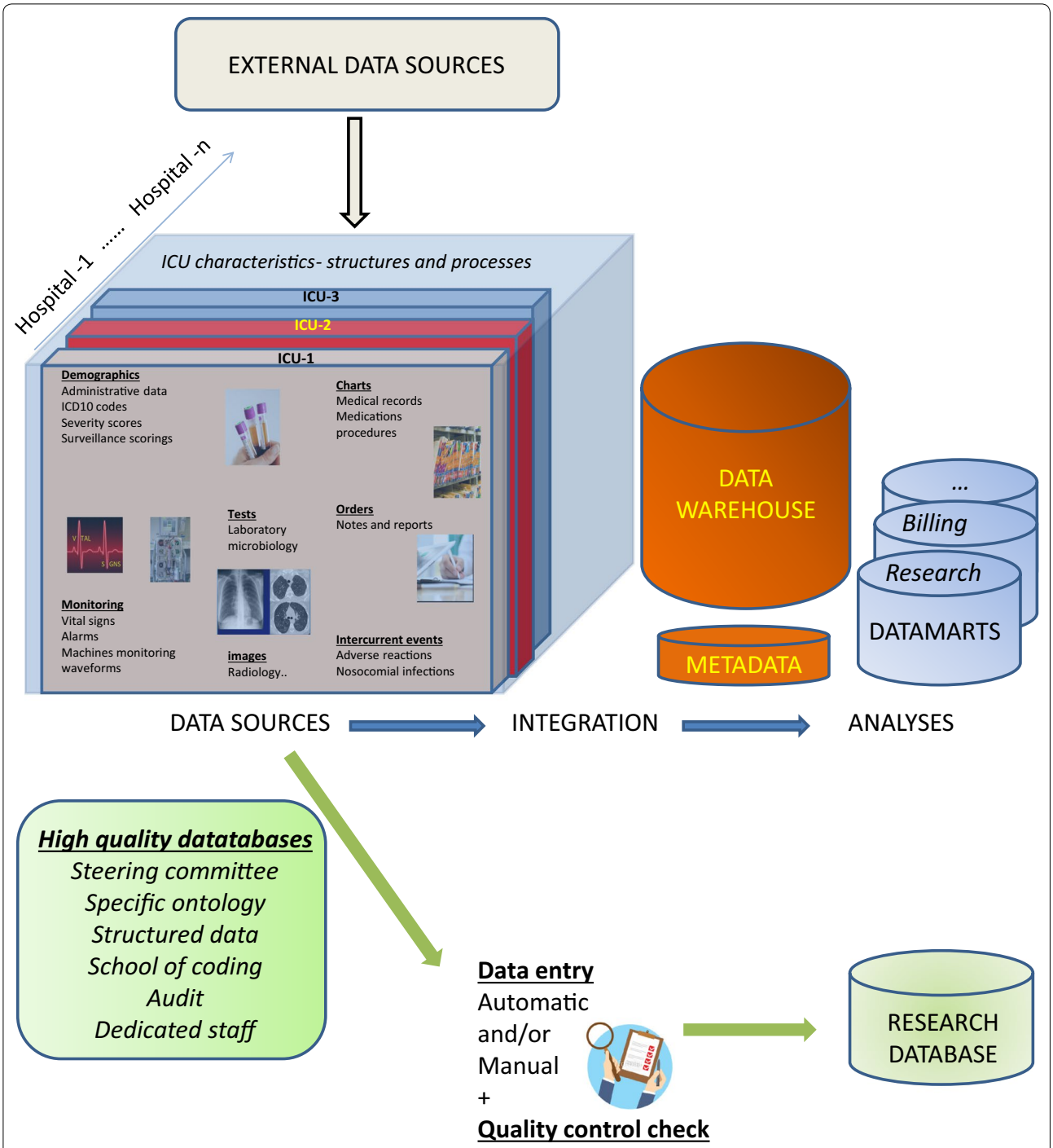

Fig. 1 From data warehouse to high-quality databases for research. Data sources are now electronically recorded. This considerable amount of data with various structured or unstructured formats and ontologies needs to be integrated into data warehouse using metadata. Metadata is defined as the data providing information about one or more aspects of the data. After integration, these data are organized into data marts easily accessible for many purposes such as research. In this organization structure, the data are not captured for research. Another way to use a data source is to integrate it into high-quality databases devoted to research purposes. This organization directly provides specific variables with precise definition and temporality but requires dedicated personnel, specific training, and careful organization. Both organization structures possess different advantages and pitfalls and may be used as helpful research tools

Author details
[1] Medicine and Infectious Diseases ICU (MI2), APHP Bichat Hospital, Paris, France. [2] IAME U 1137 Inserm-Paris Diderot University, Paris, France. [3] Intensive Care Unit, Delafontaine Hospital, 93205 Saint Denis, France. [4] Department of Biostatistics and Clinical Research, CHU de Caen, Caen, France. [5] EA2656 Groupe de Recherche sur l'Adaptation Microbienne (GRAM 2.0), Caen Normandie Université, Caen, France.

Compliance with ethical standards

Conflicts of interest
JFT is the director of the OUTCOMEREA network. The authors have no other conflict of interest to declare.

Received: 20 October 2018 Accepted: 24 October 2018
Published online: 14 December 2018

References

1. Prasad V, Vandross A, Toomey C, Cheung M, Rho J, Quinn S, Chacko SJ, Borkar D, Gall V, Selvaraj S, Ho N, Cifu A (2013) A decade of reversal: an analysis of 146 contradicted medical practices. Mayo Clin Proc 88:790–798
2. Celi LA, Mark RG, Stone DJ, Montgomery RA (2013) "Big data" in the intensive care unit. Closing the data loop. Am J Respir Crit Care Med 187:1157–1160
3. Frieden TR (2017) Evidence for health decision making—beyond randomized, controlled trials. N Engl J Med 377:465–475
4. Rothwell PM (2005) External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet 365:82–93
5. Pocock SJ, Stone GW (2016) The primary outcome fails—what next? N Engl J Med 375:861–870
6. McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C (1999) Interpreting the evidence: choosing between randomised and non-randomised studies. BMJ 319:312–315
7. Kitsios GD, Dahabreh IJ, Callahan S, Paulus JK, Campagna AC, Dargin JM (2015) Can we trust observational studies using propensity scores in the critical care literature? A systematic comparison with randomized clinical trials. Crit Care Med 43:1870–1879
8. de Grooth HJ, Parienti JJ, Oudemans-van Straaten HM (2018) Should we rely on trials with disease- rather than patient-oriented endpoints? Intensive Care Med 44:464–466
9. Bailly S, Meyfroidt G, Timsit JF (2018) What's new in ICU in 2050: big data and machine learning. Intensive Care Med 44:1524–1527
10. Simpkin AL, Schwartzstein RM (2016) Tolerating uncertainty—the next medical revolution? N Engl J Med 375:1713–1715
11. Myburgh J, Cooper DJ, Finfer S, Bellomo R, Norton R, Bishop N, Kai Lo S, Vallance S (2007) Saline or albumin for fluid resuscitation in patients with traumatic brain injury. N Engl J Med 357:874–884
12. Guidet B, Aegerter P, Gauzit R, Meshaka P, Dreyfuss D (2005) Incidence and impact of organ dysfunctions associated with sepsis. Chest 127:942–951
13. Truche AS, Darmon M, Timsit JF (2016) Renal replacement therapy modalities in the ICU: the continuity is intermittent—response to comments by Schefold. Intensive Care Med 42:1840–1841
14. Biondi-Zoccai G, Romagnoli E, Agostoni P, Capodanno D, Castagno D, D'Ascenzo F, Sangiorgi G, Modena MG (2011) Are propensity scores really superior to standard multivariable analysis? Contemp Clin Trials 32:731–740
15. Austin PC (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivar Behav Res 46:399–424
16. Shintani AK, Girard TD, Eden SK, Arbogast PG, Moons KG, Ely EW (2009) Immortal time bias in critical care research: application of time-varying Cox regression for observational cohort studies. Crit Care Med 37:2939–2945
17. Bailly S, Pirracchio R, Timsit JF (2016) What's new in the quantification of causal effects from longitudinal cohort studies: a brief introduction to marginal structural models for intensivists. Intensive Care Med 42:576–579
18. Mazumdar M, Messinger S, Finkelstein DM, Goldberg JD, Lindsell CJ, Morton SC, Pollock BH, Rahbar MH, Welty LJ, Parker RA (2015) Evaluating academic scientists collaborating in team-based research: a proposed framework. Acad Med 90:1302–1308