VALUE IN HEALTH 20 (2017) 1023–1024

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Making Real-World Evidence More Useful for Decision Making

Real-world evidence (RWE) holds enormous promise, with some of that promise beginning to be realized in the evaluation of harms. However, in order to accomplish major strides in harms assessment, and ultimately in the evaluation of effectiveness, many steps have to be taken. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and International Society for Pharmacoepidemiology (ISPE) papers [1,2] outline those steps needed to have observational studies, based on data routinely collected in practice, "more closely approximate" randomized controlled trials (RCTs) [3]. The goal is praiseworthy because of the aspiration to be able to draw "causal conclusions" in combination with experimental studies (randomized controlled trials, RCTs), considered the gold standard for use as a major source of evidence, where data is allocated by investigators. However, RCT data are also increasingly being seen as problematic not only because of the costs, the length of studies, and the exclusion of major segments of the population, but also because of the controversies surrounding RCT interpretation. All of these factors conspire to make RCTs (even pragmatic clinical trials, PCTs) not as valuable as once believed, and cannot be regarded as the only mode of providing data for medical decision making. Thus, there is a need to have observational studies based on real-world data to come up with findings that can lead to trustworthy clinical practice guidelines and decision aids.

The ISPOR paper lays out the potentially solvable barriers in its general recommendations [1]. The recommendations are parsimonious and refer to the companion paper developed by ISPE exploring the details involved in creating the transparency in the conduct of RWE studies needed for reproducibility [2]. Reproducibility is necessary for all research designs, but is particularly important for observational studies where relatively little is under the control of investigators.

The first two recommendations are related: post the protocol and analysis plan on a public registration site, but only for confirmatory studies with testable hypotheses. Not only does the posting allow systematic reviewers to capture the potential universe of studies, but it also allows the public to examine whether core issues are being addressed: the stating of the hypothesis, the formation of a control group, the identification of variables and covariates to form key subgroups, and agreement about the outcomes used.

The third recommendation addresses publication, focusing on any deviation from the original intent of the study. The fourth recommendation calls for full transparency, including data sharing such that the data can be reused. The fifth recommendation emphasizes the value of confirmation of the results in a second data source. Much as regulatory agencies typically require more than one RCT to establish credible evidence, more than one HETE (Hypothesis Evaluating Treatment Effectiveness) observational study examining different patient populations typically should be required by regulatory authorities to be considered credible evidence. The sixth recommendation is to publicly address methodologic criticism of a study following publication. Finally, the seventh recommendation focuses on stakeholder involvement with the appropriate caution, due to lack of evidentiary support, on the value of specific stakeholder roles.

Implementing the recommendations for the procedural practices outlined in the report, as well as the methodological considerations, is a tall order and is seemingly an unrealistic one in the current climate. However, two considerations may provide impetus to the goals of the report. One, as described earlier, is the now well-appreciated problem of basing treatment decisions on randomized trial data alone. The other propelling force, as outlined in the recent commentary by Jarow et al. is the possibility of using hybrid data sources, complementing the database with drill-down information from electronic medical records and even smart phones [4]. In a hybrid study, the EMR could supply information that would confirm diagnoses, imaging, genetic testing and results, and even medication regimens. Smart phones could supply patient reported information on social determinants of outcomes, on behaviors such as smoking, on comorbidity and on mental health [5]. The comprehensiveness and richness of these variables, derived from the EMR, from patient input, and from other sources, could move observational and real-world data closer to randomized trials, while retaining all of the advantages of large observational data. Indeed, this is the direction that observational studies are moving toward.

Two recent types of studies may illustrate this notion of combining complementary data sources. One is the seeming never-ending controversy over whether azithromycin "causes" heart disease and leads to cardiovascular mortality [6]. A number of well-performed real-world data studies have been executed, a tour de force because literally millions of patients have been studied to examine this relationship, with some saying yes and some saying no. One of the more recent studies titled ARITMO (Arrhythmogenic Potential of Drugs), published last year in *Canadian Medical Association Journal* from the European consortia, explored the relationship between taking azithromycin and ventricular arrhythmias but included a "manual review" of a random sample of medical records to adjudicate the outcome [7]. In addition, in one of the national databases, identified cases were validated by manually examining the medical records as part of a larger initiative in harmonizing data extraction [8]. Multiple converging sub-studies from the same populations, or independent studies combining multiple data sources, could bring real-world data closer to "causality" and could be perceived as acceptable alternatives to randomized trials. Indeed, as the mining of free-text data becomes more accurate through natural language processing and machine learning approaches, such analyses will be more readily executed.

Another example is the Comparative Effectiveness analysis of Surgery and Radiation (CEASAR) study, a national observational

study of early prostate cancer treatment across multiple institutions and data bases (disclosure: I was co-investigator on the study) [9,10]. This was not a large "health care database" study, but it identified over 2500 patients in the SEER databases nationally. It used SEER data for vital status at follow-up and also for supplementing missing data such as race and treatment details. Once the patients were identified, contacted and consented, they completed a questionnaire and their medical records were examined, creating a merged file on each individual patient. The addition of EMR data and patient reports provided information not available in large databases. Reliance on EMR diagnoses for case identification, for identifying co-existing diseases, for propensity score variable analyses, and for treatment indication is problematic and corroboration of diagnoses based on multiple data sources adds to the value of the study [11]. The CEASAR study is a potential model for statewide, regional, or national registries combining multiple real-world data sources in an inexpensive way. The EMR information could be standardized and the patient derived information could be obtained by automated means. The interaction of disparate data sources would maximize the potential of RWE to advance clinical decision making.

As Jarow et al. pointed out, a clear distinction between real-world and non-real-world data is a false dichotomy and, in reality, actually both "exist" on a continuum [4]. Appreciation of this notion allows for a variety of mixed data sources, independent of study designs. Attempts to follow the recommendation of ISPOR, will lead to better information on which to base decisions surrounding both safety and effectiveness. The notion promulgated by some that PCTs provide the best solution is problematic [12]. First, because PCTs suffer from many of the same problems as RCTs, and second, because the wealth of real-world data that will be mined from insights over the coming years via observational studies will be far greater in volume and cannot be ignored if we are to promote the development of a learning health care system. It will be up to stakeholders, including regulators, HTA authorities, payers, and providers to provide incentives to follow the recommendations in the ISPOR and ISPE taskforce reports. If these recommendations are followed, more rapid progress will be achieved in establishing a learning health care system.

Sheldon Greenfield, MD
Department of Medicine
University of California, Irvine
100 Theory, Suite 110
Mail Code: 5800
Irvine, CA 92697
E-mail address: sgreenfi@uci.edu

http://dx.doi.org/10.1016/j.jval.2017.08.3012

REFERENCES

[1] Berger ML, Sox H, Willke R, et al. Recommendations for Good Procedural Practices for Real World Data Studies of Treatment Effectiveness and/or Comparative Effectiveness Designed to Inform Healthcare Decisions: Report of the joint ISPOR-ISPE Special Task Force on Real World Evidence in Health Care Decision Making. Value Health 2017;20: xx-xx.
[2] Wang SV, Brown J, Gagne JJ, et al. Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies. Value Health 2017;20:1009–22.
[3] Greenfield S, Platt R. Can observational studies approximate RCTs? Value Health 2012;15:215–6. http://dx.doi.org/10.1016/j.jval.2012.01.003.
[4] Jarow JP, LaVange L, Woodcock J. Multidimensional Evidence Generation and FDA Regulatory Decision Making: Defining and Using "Real-World" Data. JAMA July 2017. http://dx.doi.org/10.1001/jama.2017.9991.
[5] Kaplan SH, Billimek J, Sorkin DH, Ngo-Metzger Q, Greenfield S. Who can respond to treatment? Identifying patient characteristics related to heterogeneity of treatment effects. Med Care 2010;48(6 Suppl):S9–16. http://dx.doi.org/10.1097/MLR.0b013e3181d99161.
[6] Ray WA, Murray KT, Hall K, Arbogast PG, Stein CM. Azithromycin and the risk of cardiovascular death. N Engl J Med 2012;366, May 17, 2012.
[7] Trifirò G, Ridder M de, Sultana J, et al. Use of azithromycin and risk of ventricular arrhythmia. Can Med Assoc J 2017;189:E560–8. http://dx.doi.org/10.1503/cmaj.160355.
[8] Avillach P, Coloma PM, Gini R, et al. Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. J Am Med Inform Assoc JAMIA 2013;20:184–92. http://dx.doi.org/10.1136/amiajnl-2012-000933.
[9] Barocas DA, Alvarez J, Resnick MJ, et al. Association between radiation therapy, surgery, or observation for localized prostate cancer and patient-reported outcomes after 3 years. JAMA 2017;317:1126–40. http://dx.doi.org/10.1001/jama.2017.1704.
[10] Barocas DA, Chen V, Cooperberg M, et al. Using a population-based observational cohort study to address difficult comparative effectiveness research questions: the CEASAR study. J Comp Eff Res 2013;2:445–60. http://dx.doi.org/10.2217/cer.13.34.
[11] Daskivich TJ, Abedi G, Kaplan SH, Skarecky D, Ahlering TE, Greenfield S. Electronic Medical Record Problem List: Accurate Enough for Risk Adjustment? *Am J Manag Care* (In press).
[12] Sox HC. Pragmatic Trials: Practical Answers to "Real World" Questions. JAMA 2016;316:1205–6. http://dx.doi.org/10.1001/jama.2016.11409.