Using observational data for personalized medicine when clinical trial evidence is limited

Boris Gershman, M.D.,[a,b] David P. Guo, M.D.,[a,b] and Issa J. Dahabreh, M.D., M.S.[c,d,e]

[a] Division of Urology, Rhode Island Hospital and the Miriam Hospital; [b] Warren Alpert Medical School, [c] Center for Evidence Synthesis in Health, School of Public Health, [d] Department of Health Services, Policy and Practice, School of Public Health, and [e] Department of Epidemiology, School of Public Health, Brown University, Providence, Rhode Island

Randomized clinical trials are considered the preferred approach for comparing the effects of treatments, yet data from high-quality clinical trials are often unavailable and many clinical decisions are made on the basis of evidence from observational studies. Using clinical examples about the management of infertility, we discuss how we can use observational data from large and information-rich health-care databases combined with modern epidemiological and statistical methods to learn about the effects of interventions when clinical trial evidence is unavailable or not applicable to the clinically relevant target population. When trial evidence is unavailable, we can conduct observational analyses emulating the hypothetical pragmatic target trials that would address the clinical questions of interest. When trial evidence is available but not applicable to the clinically relevant target population, we can transport inferences from trial participants to the target population using the trial data and a sample of observational data from the target population. Clinical trial emulations and transportability analyses can be coupled with methods for examining heterogeneity of treatment effects, providing a path toward personalized medicine. (Fertil Steril® 2018;109:946–51. ©2018 by American Society for Reproductive Medicine.)
Key Words: Causal inference, emulating clinical trials using observational data, heterogeneity of treatment effects, study design, transportability

BACKGROUND

Randomized clinical trials are considered the preferred approach for comparing the effects of treatments because randomization renders the compared groups similar (in expectation) with respect to both measured and unmeasured (including unknown) pretreatment covariates, and justifies the use of straightforward statistical methods to estimate treatment effects (1). Clinical trials are prospectively planned experimental studies; thus, besides randomization, they have many other features that enhance validity, such as concurrent control groups, standardized outcome definitions and follow-up procedures, and measures to limit missing data and loss to follow-up. For these reasons, traditional "evidence hierarchies" identify clinical trials or meta-analyses of clinical trials as level I evidence (the highest possible) (2–4).

Despite the recognition that well-conducted clinical trials can support valid causal inference, physicians often have to make clinical recommendations with no or limited evidence from clinical trials (5, 6). Clinical trials are often infeasible because of logistical, cost, or ethical considerations (7–10). And when conducted, clinical trials sometimes suffer from serious methodological shortcomings, such as selection bias (e.g., informative dropout and loss to follow-up) and missing data, or have sample sizes and follow-up durations that are inadequate for assessing comparative effectiveness for clinically important outcomes (3, 5, 6, 11). Even in high-quality clinical trials, trial participants are often selected on the basis of characteristics that modify the treatment effect. When that is the case, estimates of population-averaged treatment effects from trial participants do not directly apply to the patient populations seen in clinical practice.

In this article, using two clinical cases, we discuss how we can use observational data from large and information-rich health-care databases combined with modern epidemiological and statistical methods to draw inferences about the effects of treatments

Received March 3, 2018; revised March 28, 2018; accepted April 3, 2018.
B.G. has nothing to disclose. D.P.G. has nothing to disclose. I.J.D. has nothing to disclose.
B.G. and D.P.G. should be considered similar in author order.
Supported in part through two Patient-Centered Outcomes Research Institute (PCORI) Methods Research Awards (ME-1306–03758 and ME-1502–27794 to I.J. Dahabreh). All statements in this paper, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of the PCORI, its Board of Governors, or the Methodology Committee.
Reprint requests: Boris Gershman, M.D., Rhode Island Hospital and The Miriam Hospital, Warren Alpert Medical School of Brown University; 195 Collyer Street, Suite 201; Providence, Rhode Island 02904 (E-mail: boris.gershman@lifespan.org).

Fertility and Sterility® Vol. 109, No. 6, June 2018 0015-0282/$36.00

https://doi.org/10.1016/j.fertnstert.2018.04.005

when clinical trial evidence is unavailable or not applicable to clinically relevant target populations. When clinical trial evidence is unavailable, we can conduct observational analyses emulating a hypothetical pragmatic target trial that would address the clinical question of interest. When clinical trial evidence is available but not applicable to the target population, we can transport inferences from trial participants to the target population using the trial data and a sample of observational data from the target population. Both trial emulation and transportability analyses can be combined with methods for examining the heterogeneity of treatment effects to personalize care.

USING OBSERVATIONAL DATA TO EMULATE TARGET TRIALS

Clinical Case 1

A 32-year-old man with primary infertility presents with his wife of the same age. He is diagnosed with nonobstructive azoospermia, and their physician recommends surgical sperm retrieval with in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI). The couple asks whether fresh or cryopreserved sperm would increase the chance of a clinical pregnancy. What evidence should the physician rely on to counsel the couple?

Methodological Considerations for Clinical Case 1

For the couple in this case, no clinical trials have compared IVF-ICSI with fresh versus cryopreserved surgically retrieved sperm; a recent systematic review on this question identified 11 observational studies but no randomized trials (12). In the absence of clinical trial evidence, well-conducted observational studies are often the best source of evidence (3, 5, 11).

Observational studies take advantage of clinical practice variation to assess the effects of treatments that are not assigned by the investigators. Because treatment assignment is not randomized, observational studies are susceptible to confounding bias by shared causes of the treatment and the outcome. For instance, physicians may be more likely to offer cryopreservation to men with a higher probability of having viable sperm; and the availability of treatment may vary by geographic location or socioeconomic status, which may also affect fertility rates. In addition, observational studies are susceptible to selection bias (like any follow-up study, including clinical trials), measurement error bias, and (when considering per-protocol effects) time-varying confounding (i.e., when the exposure is time-varying and a covariate measured after the baseline is an independent predictor of both subsequent treatment and the outcome, within strata determined by baseline covariates and prior treatment [13–15]). Because design choices that can mitigate selection and measurement error biases are often impossible to implement in observational studies (especially when using routinely collected data), observational analyses may be more susceptible to these biases than clinical trials. Thus, causal inference from observational studies is often more speculative than inference based on well-conducted clinical trials, and the conduct of observational studies needs great care.

When designing an observational study, it is useful to consider a *hypothetical target trial* that would address the same clinical question (14, 16, 17). The process begins by specifying the protocol of this target trial: eligibility criteria, treatment strategies, assignment procedures, follow-up duration, outcomes, causal contrasts (i.e., targets of inference, such as the intention-to-treat effect), and analysis plan (17). The protocol is used to guide the conduct of the observational study in an iterative process: refinements to the clinical question and practicalities related to the data suggest modifications of the protocol, while keeping the target trial in view ensures that the data used for the emulation contain adequate information and are processed in a way that can allow a causal interpretation of the final analysis. In the context of using routinely collected clinical (e.g., electronic medical records) or administrative (e.g., insurance claims) data (17), the target trial framework provides a way to impose structure on messy "big data" and "real-world evidence." Given that data collection occurs as part of regular care encounters primarily for nonresearch purposes, the target trials that can be emulated with routinely collected data are necessarily highly pragmatic ones. For example, the target trials would define interventions fairly broadly, use administrative data to ascertain outcomes, and forego blinding (18, 19).

The intuition that an observational study comparing treatments should be viewed as an attempt to emulate a target trial is shared across different fields that have to rely on observational analyses to compare treatments, including medicine, epidemiology, and the social sciences (20–32). This intuition has motivated various "benchmarking" attempts comparing estimates from observational studies against matched clinical trials (20-32). In medicine, these comparisons have shown that good agreement between observational studies and clinical trials is possible, but fairly large disagreements do occur, even if they are rarely statistically significant (5). Most comparisons have relied on matching already completed observational and randomized studies conducted independently in different patient populations, using the incomplete information available in the published literature (as opposed to patient-level data and study protocols), without harmonizing the methods for baseline confounding control (in the observational studies) or addressing selection and measurement error bias (in either design). Observational studies designed to explicitly emulate target trials combined with better data and state-of-the science methods to address methodological shortcomings (in both clinical trials and observational analyses) should lead to better agreement.

The target trial framework encourages clear thinking about the goals of the observational analysis and ensures that the methods employed can fit those goals. Thus, the framework confers many practical benefits, some of which we have collected in Table 1. Practical experience with observational analyses explicitly designed to emulate target trials is relatively limited. However, initial results from diverse fields of application are promising (34–37). Much of the value of the target trial framework derives from discussions