The Value of Statistical Analysis Plans in Observational Research
Defining High-Quality Research From the Start

Laine Thomas, PhD
Eric D. Peterson, MD, MPH

The increasing availability of electronic health data combined with federal investment has stimulated an expansion in observational clinical research.1 Observational studies can complement clinical trials and provide important information about comparative safety and effectiveness in populations not well studied in clinical trials. However, there are numerous examples in which the findings from observational studies have failed to be replicated.2 These failures may be due to several factors, including the exploratory nature of observational questions, failure to fully account for treatment selection bias, known publication biases, and the tendency to pursue post hoc hypotheses. This later problem, termed data dredging, is facilitated by the lack of fidelity to a prespecified hypothesis and inadequate reporting of the actual analytic process.

In contrast to observational research, clinical trials ordinarily operate under strict standards at every step of study planning and data analysis. A detailed protocol, including the definition of end points, hypotheses, and all analytical procedures, is submitted to the US Food and Drug Administration and registered in various data repositories, such as clinicaltrials.gov, prior to enrollment of patients. Trial registration helps ensure that both positive and negative findings are publicly known. Prespecification of trial protocols creates an incentive to understand the biological function of the intervention, carefully define the population of interest, target the most appropriate end points, and achieve certainty about the statistical approach. Prespecification of hypotheses and minimal testing means that standard errors and P values are accurate measures of uncertainty and statistical evidence is rigorous. Trial protocols can also be referred to and reviewed to understand the questions, end points, and subgroup analyses that were defined ahead of time and those that were post hoc and in need of replication for validation.

A natural question arises as to whether elements of this rigorous process should be applied to observational research. While select observational studies are already registered in clinicaltrials.gov,3 some argue that observational research is, by its nature, exploratory and requires substantial flexibility to investigate novel findings and unexpected signals in the data.4,5 Yet interpretation of statistical evidence (P values and confidence intervals) can be made potentially meaningless when multiple hypotheses are generated by exploring the available data. Hence, a balance must be achieved to promote some flexibility but also encourage a rigorous, efficient analytical process that minimizes unnecessary data dredging.

Aside from considering the advantages of preregistration, substantial progress has been made to define standards for reporting observational research.6,7 The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations provide a checklist of items to include in reporting of cohort, case-control, and cross-sectional studies.6 Good Research for Comparative Effectiveness (GRACE) principles similarly reflect a consensus on good practice for design and evaluation in comparative effectiveness research.7 Despite these standards for high-quality observational research and reporting, such guidelines are not consistently adopted, in part because of their complexity and the difficulty of including all components in published articles.

The concepts for improving observational research can be operationalized via use of a formal, prospectively defined statistical analysis plan (SAP). The SAP should include enough detail that another statistician familiar with the data set (or their own independent data) could replicate the analysis. This implies that the SAP should delineate populations (exclusion criteria); end points; descriptive objectives; testable hypotheses; modifications or derivations of standard variables; statistical methods, including handling of missing data, correlated data, bias, and confounding; subgroups; interactions; and sensitivity analy-
While developing the SAP, the statistician may investigate the project and anticipate issues. The stated hypotheses help avoid misunderstandings between cross-checking of potential conclusions with technically the aims and approach. The review of table shells and the corresponding potential conclusions or takeaway details mentioned above, table shells for intended output, proposed to address each major objective, including the objectives and testable hypotheses. Statistical methods are SAP, translating the clinical questions into descriptive publications committee. From the proposal, a primary statistical the gains in public and academic trust associated with transparency outweigh this concern.9 Thus, investigators conducting observational research should develop and use prespecified SAPs and should submit these to journals, along with their manuscripts, for review and ultimate online publication.

**Conflict of Interest Disclosures:** The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

**REFERENCES**

5. Pearce N. Registration of protocols for observational research is unnecessary and would do more harm than good. Occup Environ Med. 2011;68(2):86-88.
9. Ioannidis JPA. The importance of potential studies that have not existed and registration of observational data sets. JAMA. 2012;308(6):579-576.