Editor’s Perspective

Registries and Selection Bias
The Need for Accountability

Harlan M. Krumholz, MD, SM

Outcomes researchers share a strong interest in determining what is being achieved in practice. In the quest to reveal what is actually being done in medicine, we are often in a position to use existing data about clinical practice and patient experience to answer relevant questions, rather than having to generate data de novo. Patient registries, with their size, scope, and focus on patients who represent those seen in clinical practice, have become an increasingly common source of data for outcomes researchers.

There is much to like about patient registries. These organized collections of data about clinical practice are a vehicle by which we can assemble massive amounts of information about patients and use it to assess quality of care, identify patterns of care, facilitate real-world research, and improve performance. In cardiovascular medicine, registries capture information about patients with acute coronary syndromes and heart failure as well as those who undergo procedures related to cardiac catheterization, percutaneous coronary intervention, carotid stents, and internal cardioverter defibrillators. Federal agencies are turning to registries to provide information about quality and produce insights about how to best practice medicine through an emphasis on comparative effectiveness research.

Outcomes research articles based on registry data are numerous, including those from NRMI, CRUSADE, GWTG-HF, GWTG-Stroke, NCDR ACTION, NCDR Cath PCI, NCDR ICD, and others. In our first year, Circulation: Cardiovascular Quality and Outcomes published several high-profile articles based on registries.1–4

Amid the enthusiasm for registries, Ferreira-Gonzalez et al5 publish an article in this issue of the journal that sends up a warning flare about the quality of registry data. These investigators used MASCARA, a Spanish registry of patients with an acute coronary syndrome, to investigate how well individual hospitals were able to comply with the instructions to enroll consecutive patients. Of the 50 hospitals participating in MASCARA, only 17 agreed to undergo scrutiny about how well they followed the enrollment directions of the registry. These 17 hospitals had enrolled 3265 patients in the registry. The quality audit revealed that, in the same time period, another 1439 eligible patients were not enrolled. Compared with enrolled patients, these patients had higher risk and received poorer quality of care, and had a 3-fold higher in-hospital mortality.

This finding is notable for many reasons. A number of registries delegate responsibility for patient selection and data collection to sites and individuals without experience and training in research methods. They are commonly operated as part of the clinical quality activities of an organization, rather than as observational studies. As a result, many individuals with registry-related responsibilities may also have a considerable amount of unrelated responsibility that impacts their time and ability to enroll all eligible patients. Oversight of such data collection efforts may also be minimal because of cost considerations. Consequently, there has been some concern about data quality and some registries have implemented robust quality assurance activities. These activities, however, have focused predominately on data accuracy and reliability.

Although data accuracy and reliability are important to the registry, proper patient selection is at least as important. Ferreira-Gonzalez et al demonstrate that a substantial proportion of eligible patients for MASCARA were not enrolled and that those patients had a markedly different profile, biasing the cohort toward a healthier group that received higher quality care. Our concern about the generalizability of clinical trial populations that often do not enroll typical patients and seem to systematically exclude those with important age, sex, race, and comorbidity characteristics is not considered to be a problem with registries. In fact, a frequently touted strength of registries is their ability to portray the experience of typical patients—individuals who were not screened out because of trial-specific inclusion or exclusion criteria. The patient selection bias that is illuminated in this study raises a troubling concern that registries—even those designed to enroll consecutive patients—are not immune to generalizability issues.

The consequences of selection bias on the results of such registry-based research are profound, including risk prediction scores that misrepresent risk in patients who are excluded, or selection of risk factors that do not apply to the population broadly. Virtually any research question might be affected in ways that are hard to predict.

The findings also have important implications for quality assessment. The disparity in quality of care between those selected and not selected indicates the potential for misrepresentation of hospital performance. Ironically, if there is variability in adherence to the registry design, institutions that
adhere more closely to the rules may appear to perform worse.

There may even be inherent issues in the methodology used by Ferreira-Gonzalez et al. Although for many conditions there is a degree of uncertainty, leading to disagreement, about whether some patients meet criteria, the pertinent question that remains is whether the instructions are being followed in a way that is reproducible. The magnitude of the discrepancy between the registry enrollment at the hospitals and the quality audit conducted by Ferreira-Gonzalez and colleagues suggests that this is more than a minor disagreement.

How widespread is the problem? Although it is conceivable that the issues with MASCARA were peculiar to that registry, the mere possibility that selection bias is widespread among registries should initiate immediate efforts to assess the extent of the problem. Because every registry is unique in its implementation, variability in the selection bias among registries is likely. As we embark on a research era in which “real-world” registries are assuming an increasingly prominent role, understanding the prevalence and magnitude of such potential selection biases is critically important.

Registries have made the leap from second-tier status in American medicine, used locally by institutions that collected data for the purpose of improving care, to their current place at center stage. Registry data are poised to serve as an instrumental means of producing knowledge that will guide practice, as well as a source of information about relative performance, with potential revenue consequences. However, with the increasing prominence and utility comes increasing responsibility. In the future, registries will need to demonstrate their performance with respect to producing high-quality data. Success in this effort is critical to the achievement of their immense promise.

Disclosures
None.

References

Key Words: outcomes research ■ registries ■ selection bias
Registries and Selection Bias: The Need for Accountability
Harlan M. Krumholz

doi: 10.1161/CIRCOUCOMES.109.916601
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 American Heart Association. Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/2/6/517

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org/subscriptions/