Reliability of Medicare Claim Forms for Outcome Studies in Kidney Transplant Recipients: Epidemiology in Clinical Outcome Trials

Lisa Nanovic* and Bruce Kaplan†

*Department of Medicine, Section of Nephrology, University of Wisconsin Hospital and Clinics, Madison, Wisconsin; and †Department of Medicine, Pharmacology and Surgery, Section of Nephrology, Arizona Health Science Center, University of Arizona School of Medicine, Tucson, Arizona

Clin J Am Soc Nephrol 4: 1156–1158, 2009. doi: 10.2215/CJN.03300509

Medical databases may offer substantial opportunities for outcomes research to investigators in the field of renal transplantation. An increasing number of analyses have attempted to merge Medicare billing claim forms with large, national registries. Large registries such as the US Renal Data System and the Scientific Registry of Transplant Recipients have afforded researchers the opportunity for large-scale, population-based analysis; however, these registries have come under increasing scrutiny in terms of their use for associative outcome studies. It has been postulated that by merging Medicare claims with other large registries, some of the limitations outlined by critics can be overcome. Despite the increasing number of publications that are taking this approach, this practice has not been fully validated in the renal transplantation population.

Four main elements that characterize a medical database are (*1*) population, (2) medical events, (*3*) coding systems, and (*4*) data management (1). Each of the registries mentioned in the previous paragraph, to some degree, do not capture all of the essential elements. This may introduce nonrandom and uncaptured bias, potentially leading to nonindependent associations. In addition, the large amount of variables available can lead to type 1 error if the association found is not robust across a number of situations and does not have some biologic plausibility.

In population-based registries, inadequate capture of certain data is unavoidable (2). Eligibility criteria may be difficult to obtain properly, with imprecise and inconsistent definitions limiting the accuracy of the data used (3). Also, only selected medical events are included in databases. In administrative and disease databases, only demographic data, along with what is required for financial allowance in the former and information regarding the target disease in the latter, may be available. Coding systems (*e.g.*, those used in the Medicare claim form database) have their own drawbacks, including inconsistency in the detail of coding. In *International Classification of Diseases, Ninth Revision* (ICD-9), the specificity of diagnoses is often more limited than that for procedures, and ICD-9 codes allow diagnoses to be classified as "rule out" rather than a confirmed disease entity (4,5). Furthermore, coding definitions may be vague, which can lead to subjective and often inconsistent designations of disease or disease processes (6). In addition, it is often not clear whether a diagnosis or treatment is new (7) or whether a coded diagnosis is a comorbidity or a complication. Changes in the coding systems may introduce a lack of comparability between previous and new identifiers (8). As with any database, human error must also be recognized. This potential for human error can be partially attributed to thousands of individuals who enter data (9–12), introducing potential error and variation (6,13). Errors may arise from the coders, the coding systems, or the medical chart itself (13). The level of importance of information to be coded, as well as reasoning behind the coding (*e.g.*, reimbursement), is at the perception of the individual who enters the data.

Critical to using any large database is careful consideration of the question being asked and the structure of that medical database. The aforementioned issues can create daunting statistical issues. The large sample sizes in database research can create multiple comparisons, leading to chance associations (14). Confounding may be a difficult issue in database research because desirable covariate information may not be available (15). Inconsistencies in inclusion criteria in databases, as well as variability in coding practices, can create selection bias. It must be recognized that use of multiple data repositories enhances the reliability of the data only to the degree by which it consistently adds data, decreasing missing nonrandom confounders.

Even with known disadvantages and limitations for use of medical records for research, the merging of separate files or databases may provide a benefit. Assimilation of data sources creates an integration of different coding systems, necessitating algorithms to define cases and resolve conflicting data. Potential ideas of enhancing database research

Published online ahead of print. Publication date available at www.cjasn.org.

Correspondence: Dr. Lisa Nanovic, Home Dialysis Program, University of Wisconsin Madison Medical School, 3034 Fish Hatchery Road, Suite B, Madison, WI 53713. Phone: 608-270-6563; Fax: 608-270-5677; E-mail: ln3@medicine.wisc.edu

ISSN: 1555-9041/407–1156

through linkage of several databases have been posed to improve accuracy (2). In this issue of *CJASN*, Lentine *et al.* (16) examine the accuracy of Medicare claim forms for the diagnosis of cardiovascular events in kidney transplant recipients. The authors used Medicare Parts A and B claim forms, linking these claims to two independent databases, to evaluate the accuracy of claims representing clinical diagnoses of cardiovascular events under five different algorithms. This effort was done in a technically superb manner and addressed an important question to those involved in renal transplant outcome research.

Lentine *et al.* in this issue of *CJASN* address many of the concerns related to large databases. The authors attempt to validate the accuracy of Medicare claim forms as it pertains to cardiovascular events in the kidney transplant population. The electronic medical records from the Washington University Kidney Transplant Program Database were linked at the patient level with the national Organ Procurement and Transplantation Network records and Medicare billing claims as compiled within the US Renal Data System. They focused their measure of clinical events to cardiovascular diagnoses and procedures. Coronary heart disease, cerebrovascular disease, peripheral vascular disease, congestive heart failure, atrial fibrillation, and venous thromboembolism were the specific diagnoses, and cardiac catheterization, coronary artery bypass grafting, amputation, and revascularization of peripheral vascular disease were the specific procedures evaluated.

What the authors found was the combination of both physician and institution claims (Medicare A and B), rather than taken individually, increased the sensitivity of claims reflecting clinical diagnoses. Within a 30-d window of claims matching clinical cardiovascular events, there was perfect capture of procedures, with reduced sensitivity in claims that matched cardiovascular diagnoses, even when compared with an extended 90-d window for claims. This was independent of whether the claims were evaluated separately or together.

Lentine *et al.* found in a unique patient population an algorithm that provides sensitivity >90% to match billing claims of kidney transplant recipients to clinical cardiovascular events, allowing for a potential, accurate, data set for epidemiologic research. As pointed out in the limitations for their study, it must be recognized that the national data registries used were linked to a local hospital. This makes the sensitivity unique to this local hospital, as other hospitals may indeed have disparities in their population, description and documentation of medical events, coding, and data management on comparison. In addition, it must be kept in mind that although procedures were captured with high sensitivity, diagnoses were not. Because many important clinical events are not tied to procedures, one must wonder under which circumstances claim forms enhance associations and under which circumstances may they only add more confounders without enhancing reliability or reproducibility?

A final important point discussed by the authors was the possibility that the increased sensitivity of the combined sources could result in decreased specificity; however, additional data may potentially enhance specificity by adding appropriate exclusionary information. In using any database, an understanding of the accuracy of the data fields will enable investigators to avoid weak information in favor of more valid data points. There is no doubt that ICD-9 codes can enhance outcomes research in transplantation, but this tool, as with other large databases, must be used judiciously and with full cognizance of which questions can be answered and which questions are only further complicated by this method.

Disclosures

None.

References

1. Baron JA, Weiderpass E: An introduction to epidemiological research with medical databases. *Ann Epidemiol* 10: 200–204, 2000
2. Fisher ES, Baron JA: Malenka DJ, Barrett J, Bubolz TA: Overcoming potential pitfalls in the use of Medicare data for epidemiologic research. *Am J Public Health* 80: 1487–1490, 1990
3. Lauderdale DS, Furner SE, Miles P, Goldberg J: Epidemiologic uses of Medicare data. *Am J Epidemiol* 15: 319–327, 1993
4. Iezzoni LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E, Levine PA: Coding of acute myocardial infarction: Clinical and policy implications. *Ann Intern Med* 109: 745–751, 1988
5. McMahon LF, Smits HL: Can Medicare prospective payment survive the ICD-9-CM disease classification system? *Ann Intern Med* 104: 562–566, 1986
6. Iezzoni L: Assessing quality using administrative data. *Ann Intern Med* 127: 666–674, 1997
7. Smith GS, Langolis JA, Beuchner JS: Methodological issues in using hospital discharge data to determine the incidence of hospitalized injuries. *Am J Epidemiol* 134: 1146–1158, 1991
8. Hannan EL, Racz MJ, Jollis JG, Peterson ED: Using Medicare claims data to assess provider quality for CABG surgery: Does it work well enough? *Health Serv Res* 31: 659–678, 1997
9. Lloyd SS, Rissing JP: Physician and coding errors in patient records. *JAMA* 254: 1330–1336, 1985
10. Jencks SF: Accuracy in recorded diagnoses. *JAMA* 267: 2238–2239, 1992
11. Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP: Accuracy of diagnostic coding for Medicare patients under the prospective payment system [published erratum appears in *N Engl J Med* 322: 1540, 1990]. *N Engl J Med* 318: 352–355, 1988
12. Hsia DC, Ahern CA, Ritchie BP, Moscoe LM, Krushat WM: Medicare reimbursement accuracy under the prospective payment system, 1985–1988. *JAMA* 268: 896–899, 1992
13. Green J, Wintfield N: How accurate are hospital discharge data for evaluating effectiveness of care? *Med Care* 31: 719–731, 1993
14. Romano RS, Mark DH: Bias in the coding of hospital discharge data and its implications for quality assessment. *Med Care* 32: 81–90, 1994

15. Localio AR, Hamory BH, Sharp TJ, Weaver SL, TenHave TR, Landis JR: Comparing hospital mortality in adult patients with pneumonia: A case study of statistical methods in a managed care program. *Ann Intern Med* 122: 125–132, 1995

16. Lentine KL, Schnitzler MA, Abbott KC, Bramesfeld K, Buchanan PM, Brennan DC: Sensitivity of billing claims for cardiovascular disease events among kidney transplant recipients. *Clin J Am Soc Nephrol* 4: 1213–1221, 2009

See related article, "Sensitivity of Billing Claims for Cardiovascular Disease Events among Kidney Transplant Recipients," on pages 1213–1221.