UCLA
UCLA Previously Published Works

Title

Use of Health Care Claims Data to Study Patients with Ophthalmologic Conditions

Permalink

Journal

ISSN

Authors

Stein, Joshua D
Lum, Flora
Lee, Paul P
et al.

Publication Date

DOI

Peer reviewed

Use of Health Care Claims Data to Study Patients with Ophthalmologic Conditions

Joshua D. Stein, MD, MS[1], **Flora Lum, MD**[2], **Paul P. Lee, MD, JD**[1], **William L. Rich III, MD**[3], and **Anne L. Coleman, MD, PhD**[2,4]

[1] W. K. Kellogg Eye Center, University of Michigan, Ann Arbor, Michigan.

[2] H. Dunbar Hoskins, Jr., M.D. Center for Quality Eye Care, American Academy of Ophthalmology, San Francisco, California.

[3] Northern Virginia Ophthalmology Associates, Falls Church, Virginia.

[4] Jules Stein Eye Institute, David Geffen School of Medicine and Jonathan and Karen Fielding School of Public Health, University of California at Los Angeles, Los Angeles, California.

Abstract

Objective—To describe what information is or is not included in health care claims data, provide an overview of the main advantages and limitations of performing analyses using health care claims data, and offer general guidance on how to report and interpret findings of ophthalmology-related claims data analyses.

Design—Systematic review.

Participants—Not applicable.

Methods—A literature review and synthesis of methods for claims-based data analyses.

Main Outcome Measures—Not applicable.

Results—Some advantages of using claims data for analyses include large, diverse sample sizes, longitudinal follow-up, lack of selection bias, and potential for complex, multivariable modeling. The disadvantages include (a) the inherent limitations of claims data, such as incomplete, inaccurate, or missing data, or the lack of specific billing codes for some conditions; and (b) the inability, in some circumstances, to adequately evaluate the appropriateness of care. In general, reports of claims data analyses should include clear descriptions of the following methodological elements: the data source, the inclusion and exclusion criteria, the specific billing codes used, and the potential confounding factors incorporated in the multivariable models.

Conclusions—The use of claims data for research is expected to increase with the enhanced availability of data from Medicare and other sources. The use of claims data to evaluate resource

use and efficiency and to determine the basis for supplementary payment methods for physicians is anticipated. Thus, it will be increasingly important for eye care providers to use accurate and descriptive codes for billing. Adherence to general guidance on the reporting of claims data analyses, as outlined in this article, is important to enhance the credibility and applicability of findings. Guidance on optimal ways to conduct and report ophthalmology-related investigations using claims data will likely continue to evolve as health services researchers refine the metrics to analyze large administrative data sets.

Many studies are using claims data and other large administrative health databases to assess ophthalmology-related research questions. Findings from such studies are considered by health policymakers, third-party payers, and other decision makers as they grapple with timely challenges, such as allocating limited resources and finding ways to improve patient outcomes and patient care. Clinicians also are using data to assess their practices and performance compared with their peers and to understand the risk factors and outcomes of their patients. Other uses include estimating adherence to therapy and evaluating utilization.

Interpreting and understanding the generalizability of the findings of these types of studies can be difficult because the methodological rigor of the analyses and the extent of reporting can vary considerably. We provide an overview of the types of information contained in claims data and describe some of the advantages and limitations of using claims data for research purposes.[1] We provide a suggested checklist for authors to use in reporting analyses involving claims or other administrative health data (Table 1), as we aim to improve the overall quality and usefulness of such reports.

Information Contained in Claims Data

Numerous health care claims databases have been used by researchers. According to the International Society of Pharmacoeconomics and Outcomes Research, 382 such databases currently exist, including 153 in the United States alone (http://www.ispor.org/digestofintdb/countrylist.aspx; accessed August 29, 2013). These databases vary from large datasets capturing health care services provided to a nationwide sample of enrollees in Medicare, Medicaid, or managed-care networks, such as the Healthcare Cost and Utilization Project, to those that capture health care services in a specific community, such as Kaiser Permanente Northwest, or focus on patients with certain specific diseases, such as the Framingham Heart Study database. In some countries with a national health system, all health care encounters can be captured systematically and stored in a single repository. However, in the United States, patients can move from private plan to private plan or to a government plan, and no single source houses information on all patient encounters. United States nongovernmental databases may obtain their data through third-party insurer claims.

The types of information contained in claims databases can vary considerably. Nearly all such databases record patients' diagnosed medical conditions by using International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM),[3] billing codes and diagnostic and therapeutic procedures according to Current Procedural Terminology 4 (CPT-4)[4] billing codes. Commonly documented demographic variables include age at enrollment in the plan, sex, race, ethnicity, education, and income. Some

databases are linked to outpatient pharmacy records and results of outpatient laboratory testing; some capture the types of providers delivering the care, the site of care delivery (e.g., inpatient, outpatient, ambulatory surgical center), and the costs or charges of each service. Each data source typically contains a dictionary with a complete listing of variables.

Claims databases that have been used to study patients with ophthalmologic conditions include a 5% national sample of Medicare beneficiaries (Centers for Medicare and Medicaid Services); the Veterans Health Administration system National Patient Care Database, the Clinformatics Data Mart database (OptumInsight, Eden Prairie, MN); the PharMetrics Patient-Centric database (PharMetrics, Inc., Watertown, MA); and the Market Scan Commercial Claims and Encounter Data (Truven Health Analytics, Ann Arbor, MI). Other datasets include Medicare claims data linked with clinical, pharmacy, or survey data. Examples include the National Long-Term Care Survey, Medicare Current Beneficiary Survey, and Medical Expenditure Panel Survey. In addition, some noneU.S. administrative datasets capture care of patients in other countries.

In ophthalmology, claims data have been used to study the epidemiology of various ocular diseases[5,6]; trends in the use of eye care services, including visits to eye care providers,[7] diagnostic testing,[8–10] and medical[11] and surgical interventions[12,13]; adverse events after intraocular surgery[14–17]; associations between medical and ophthalmologic conditions[18–20]; disparities in ophthalmologic care among different populations[21]; and costs and resource consumption.[22–26]

Using Claims Data: Advantages

The potential sample sizes of studies using claims data are often larger than those of other studies. Unlike data collected from a single institution or community, these data sources often capture a relatively diverse group of enrollees receiving care in various settings across larger geographic regions.

Large sample sizes can be particularly useful for studying uncommon conditions, such as endophthalmitis. For example, 424 enrollees in one of these databases received a diagnosis of endophthalmitis in a single year, providing a potential sample size that is considerably larger than those of most other studies of endophthalmitis.

For most analyses using claims data, researchers can follow patients longitudinally to study use patterns, outcomes, and costs of care and how they change over time. Compared with most population-based cross-sectional studies, which capture the presence or absence of conditions at specific time points, claims data allow investigators to follow patients from their date of enrollment in a plan to their exit date or death. Researchers can assess the temporal relationship among different conditions, procedures, or medications with respect to one another on the basis of the date of service. For example, a researcher interested in the relationship between anti–vascular endothelial growth factor agent exposure and stroke could exclude patients who received their diagnosis of stroke before their first anti–vascular endothelial growth factor injection. Furthermore, researchers can identify patients experiencing complications months or years after surgery without much loss to follow-up

because of receipt of care by a different provider, so long as enrollment in the plan is maintained.

Claims data allow researchers to build complex multivariable models that can account for potential confounding factors. Other sources frequently provide inadequate sample sizes of people with particular conditions or outcomes of interest to permit adjustment of many confounding variables in multivariable models. Researchers should typically have at least 20 patients with the outcome of interest for each confounder added to a given regression model as per commonly recommended minimum "rules of thumb."[27] For example, if the relationship between trabeculectomy surgery and endophthalmitis were being assessed and the researchers wanted to adjust for age, sex, race (4 races), education level (4 levels), residential region (4 regions), and income (6 levels), along with comorbid diabetes, human immunodeficiency virus, and cancer, then the minimum number needed for an adequately powered analysis would be 24×20, or 480 case-patients with endophthalmitis.

Many population-based or observational studies gather information on ocular and nonocular comorbidities, prior surgeries, and medications used via interviews with the participants. Studies have shown that patient self-report can be inaccurate because many patients do not know the details about the types of medical conditions they have and the procedures they have undergone in the past.[28] Claims data on ocular and nonocular comorbidities, prescriptions, and procedures come directly from medical providers and pharmacists— sources generally considered to be more accurate than patient self-report.

Another advantage is that relationships among different conditions can be evaluated without introducing selection bias. For example, a researcher seeking to quantify the number of case-patients with endophthalmitis after retinal surgery who recruits patients from a large academic medical center may obtain an overestimate because these patients may be more complex than those seen in other practice settings. Claims data usually are not restricted to services delivered at only 1 particular medical center. Thus, findings may be more generalizable than those from other analyses.

If a researcher is interested in exploring the relationship between 2 particular interventions, the gold standard would be to conduct a randomized, controlled trial. Randomized, controlled trials allow researchers to identify causal relationships between 2 variables of interest while controlling for known and unknown confounding factors. Although a well-designed randomized trial is undoubtedly more informative than other types of study designs, including retrospective analyses using claims data, such clinical trials can be prohibitively expensive, can take years to recruit adequate numbers of patients to answer the research question of interest, and, in certain situations, may be unethical to perform. Before investing considerable resources and energy to provide a more definitive answer to a research question, researchers may find it valuable to first perform initial analyses to test their hypothesis using claims data.

Finally, all unique patient identifiers (e.g., name, address, telephone number) are removed from most claims datasets before they are made available to researchers. Therefore, these types of studies could be exempt from formal institutional review board processes.

Working with Claims Data: Issues and Pitfalls

Researchers working with claims data should be cognizant of several issues when conducting or interpreting research that uses these data sources. First, because claims data exist primarily for billing and reimbursement purposes, some of the data may incompletely capture the conditions and outcomes documented in the medical records. Patients can be misclassified because of misdiagnosis or miscoding. Miscoding may occur if a provider mistakenly submits the wrong code (e.g., codes for cataract rather than chalazion), uses a less descriptive ICD-9-CM billing code rather than a more specific code (e.g., a code of primary open-angle glaucoma rather than the specific code for uveitic glaucoma),[29] misclassifies a condition such as open-angle glaucoma when the medical record indicates normal visual field and no evidence of disease,[30] performs upcoding to maximize reimbursement (e.g., coding for a complex cataract surgery instead of a standard cataract surgery), or omits a diagnosed condition on the billing forms.[31] In one study, more than 40% of coding errors were attributable to omissions of diagnoses on the billing paperwork.[32]

A second important limitation of claims data research is that the studies are limited to only those disease entities and variables that have their own specific billing codes. Claims data usually contain no information on clinical parameters, such as best-corrected visual acuity, intraocular pressure, and findings from visual field testing or fluorescein angiography. Patients' cigarette smoking and alcohol use status, which are not captured in most claims data, cannot be studied. Moreover, with claims data, appreciating the severity of disease can be difficult; for example, patients with 1 solitary drusen and those with geographic atrophy can both be appropriately assigned the same ICD-9-CM code for nonexudative macular degeneration. Initiatives recently have been undertaken to revamp some ophthalmology-related billing codes to capture disease severity in patients with diabetic retinopathy and glaucoma (http://www.aao.org/publications/eyenet/201110/coder.cfm; accessed July 10, 2013). However, for these data to be useful for research and other purposes, providers should be educated about and encouraged to use the more detailed billing codes capturing disease severity.

When interpreting analyses using claims data, one must consider that multiple providers with different levels of experience and expertise are contributing patient data. Although some claims databases include information on the type of provider, additional information is often undocumented, such as the clinicians' board-certification status, subspecialty fellowship training, or number of years in practice. Such provider characteristics can affect the types of patients being cared for and the clinicians' ability to diagnose and treat specific conditions. Fortunately, many beneficiaries seek care from multiple different eye care providers, thus allowing an opportunity for proper disease identification with each new provider. If the same patient is assigned a code of primary open-angle glaucoma by one provider but given a more precise code of pigmentary glaucoma by other providers, researchers can consider this information before deciding which code might best reflect the patient's condition.

A third limitation is that researchers often cannot use claims data to judge the appropriateness or quality of care. Although researchers might be tempted to use claims data

to compare use rates for procedures by patient group or provider to make inferences about the provision or receipt of suboptimal care in a particular group, claims data typically contain insufficiently detailed information for such purposes.[33] Researchers interested in assessing appropriateness of care can use other approaches, for example, the Research and Development/University of California Los Angeles Appropriateness Method,[34] which is based on the consensus of experts and tested on clinical data from medical records rather than claims data. Others have advocated for "hybrid" methods that incorporate information from claims data along with medical records to judge quality and appropriateness of care.[35]

Fourth, most datasets contain insufficient detail on eye laterality to reveal whether a patient's disease is affecting 1 or both eyes. Furthermore, researchers interested in evaluating possible associations between surgical procedures and outcomes cannot know with certainty whether a documented postsurgical adverse event occurred in the eye undergoing surgery. Unlike the ICD-9-CM billing codes widely in use today, the ICD-10 billing code system requires the provider to specify which eye had the condition or procedure at issue. Thus, this limitation will probably be overcome in the near future.

Claims data research also is limited by the lack of rigorous analyses to date validating many of the existing ophthalmologic billing codes. Validation studies would demonstrate whether conditions listed from the billing encounters generally correspond to those documented in the medical record. Researchers have thus far validated billing codes for cataract and cataract surgery,[36] open-angle glaucoma,[36] diabetic retinopathy,[36] cystoid macular edema,[37] and macular degeneration[38]; however, considerably more work is needed to determine the extent to which patients with particular ophthalmologic conditions who were identified using claims data reflect those who actually have the conditions of interest according to the medical record.

In addition, claims data involve only a subset of the general population. Caution should be taken in generalizing findings to other populations. For example, results generated by using claims data from a commercial health care database may not apply to people without insurance or those residing in another country.

Claims data record only those conditions, encounters, and therapies that occur during the patients' time in the plan. Often, no historical information exists on conditions or procedures the enrollees may have had before joining the plan. For example, researchers cannot rely on claims data alone to assess the patients' duration of diabetes or whether their first diabetes-related claim reflects the timing of the initial diagnosis. Datasets that link claims data to survey data or other sources may enable researchers to identify preexisting conditions and other historical information.

The ability to perform claims-data analyses is further affected by the availability of specific billing codes. Newer technologies cannot be studied until the procedures are assigned their own CPT codes. Likewise, claims data may not distinguish between selected surgical procedures that share a single CPT code, for example, argon trabeculoplasty and selective laser trabeculoplasty. Moreover, researchers need to be aware of when particular procedural

codes became available. Before 2009, for instance, the same CPT-4 billing codes captured Descemet's stripping endothelial keratoplasty and penetrating keratoplasty.

Finally, although de-identification of claims data is important and necessary to protect patient privacy, this practice limits to some degree what researchers can study.

Reporting of Claims-Data Analyses

In this section, we describe information that researchers should generally include, typically in the Methods section, to allow proper interpretation and future comparisons with other research.

Data Source

The data sources should be adequately described. Researchers should acknowledge when the dataset documents care delivered only in particular settings (e.g., exclusively inpatient, outpatient, or skilled nursing facilities). If the claims data are linked to pharmaceutical or outpatient laboratory records, this should be stated, along with the extent to which enrollees in the medical plan are covered by the pharmacy plan.

Sample Selection

The inclusion and exclusion criteria used for the analysis should be clearly reported. This information is often easily depicted in a figure showing the process of sample selection (and corresponding patient numbers). For longitudinal studies, the authors should specify whether patients non-continuously enrolled in the plan during the study period were excluded. Including patients with noncontinuous enrollment may be problematic because care they received during interim periods between disconnected enrollment segments would be undocumented. For longitudinal analyses, often the cleanest option is to exclude patients with noncontinuous enrollment and those residing outside the United States for much of the study period. The study population should be described in detail, including whether the sample was selected from a larger, more global population (e.g., all enrollees visiting eye care providers, people enrolled during specific years, or certain age groups).

Identifying Ophthalmologic Conditions and Procedures

The specific billing codes used to identify enrollees with the diagnoses or procedures of interest should be reported. This essential information allows other investigators to replicate the analysis using another data source and permits readers to make comparisons with other studies. For example, if one study identifies patients with glaucoma using the specific ICD-9-CM billing code for primary open-angle glaucoma (365.11) and another study uses all codes for glaucoma and glaucoma suspect (365.xx) to identify patients with glaucoma, these analyses may not be directly comparable.

To elucidate the potential effects of using different groups of billing codes to identify people with different ophthalmologic conditions, Lee et al[39] compared the demographic characteristics and annual health care charges for 3 cohorts of patients with glaucoma. Each cohort was defined by specific billing codes for glaucoma: The first group was restricted to those with code 365.11; the second group was restricted to those with code 365.11 plus

codes 365.1, 365.10, 365.12, and 365.15; and the third group was restricted to those with all these codes, along with 13 others. They found differences in the demographic characteristics of enrollees among the 3 cohorts but no significant difference in total annual health care expenditures among the groups.[39] Therefore, being more or less inclusive with codes for a given analysis may or may not affect the study findings, depending on the specific study questions examined.

Investigators should carefully identify the particular patients of interest. If other researchers previously used a similar protocol to identify patients with a particular condition, the prior work can be cited. Databases vary in the number of diagnoses and procedures that can be captured at a single patient encounter. The Methods section should state how many codes are captured for an encounter and whether researchers identified enrollees with the code of interest by using only the primary diagnostic code or any code listed for the encounter. To help address concerns about misdiagnosis and miscoding of conditions, a requirement that a diagnosis be confirmed at a subsequent visit or by a separate provider can be useful.

Identifying Incident Case-Patients

Distinguishing incident from nonincident case-patients is important in analyses evaluating postsurgical outcomes. For example, in a study seeking to identify patients developing glaucoma after intravitreal corticosteroid injection, excluding patients with glaucoma documented before the injection (nonincident case-patients) would be prudent. Although researchers cannot know with certainty about the medical conditions a patient had before enrollment in the plan, techniques can be performed to help distinguish patients with a new diagnosis from others with a preexistent condition.

Researchers commonly use a "look-back" period before the index date (the date one begins following the patient for the outcome), excluding individuals with records of the condition or procedure of interest during that period. For example, to assess endophthalmitis after cataract surgery, a researcher could assign the date of the cataract surgery as the index date and exclude all individuals with an endophthalmitis diagnosis in a 1- or 2-year look-back period before this date. The longer the look-back period used, the greater the likelihood of successfully excluding nonincident case-patients. A downside to lengthening the look-back period is a reduced sample size because all patients included in the analysis need to be enrolled in the medical plan for the entire look-back period. To further ensure that identified case-patients do indeed have incident disease, researchers can require visits to eye care providers during the look-back period with no record of the diagnosis. This additional requirement can be helpful when studying conditions that are relatively asymptomatic early on. However, this may introduce bias because people who seek eye care may be more or less prone to specific ocular diseases. Sensitivity analyses can be performed to address this concern.[40]

Longitudinal Follow-up

Researchers can follow enrollees longitudinally to assess whether a particular outcome develops after an exposure. For such analyses, people with noncontinuous enrollment in the medical plan during follow-up should be excluded. Otherwise, the proportion of patients

who develop the outcome may be underestimated. For example, in an investigation of enrollees developing endophthalmitis in the 12 months immediately after cataract surgery, an enrollee who leaves the plan at 3 months after cataract extraction but does not receive her first endophthalmitis diagnosis until postsurgical month 5 would erroneously be considered to have had no infection during postsurgical follow-up. A tradeoff exists between shorter and longer follow-up periods; longer ones are more informative about risk for an outcome and how it changes over time, yet a longer follow-up usually decreases the number of patients eligible for the analysis. For longitudinal analyses, researchers should explain how they handled those with disenrollment from the plan during follow-up. An alternative strategy for handling varying lengths of follow-up is to use Kaplan–Meier plots and censor patients at their date of disenrollment or their last visit.

Assessing Medication Use

When claims data are linked to outpatient pharmacy records, researchers can study patterns of medication use and patient adherence to medical therapy. Measures to quantify medication adherence include the medication possession ratio and the proportion of days covered. For studies quantifying adherence, using a validated approach, such as one of these methods, is preferable. The Glaucoma Adherence and Persistency Study quantified glaucoma medication adherence using pharmacy claims data along with information from medical records and patient interviews and identified several problems associated with using claims data alone to study adherence: Estimates of adherence were vulnerable to errors of misidentification of newly treated patients, misclassification of added versus switched medications, and lack of knowledge about whether a patient was given medication samples.[41]

Cost Analyses

Many datasets contain information on the costs assigned to a particular patient encounter. Researchers should indicate whether the reported costs are the charges submitted by the provider to the insurance company or the actual costs paid by the carrier to the medical provider. Second, when aggregating costs over time, researchers should adjust for inflation. If truncation was performed to handle outlier costs, this too should be specified.

Multivariable Modeling

Investigators should identify and incorporate in their models all available covariates that may confound the relationship between the predictor and the outcome and inform readers of these covariates. For example, relevant confounders for outcomes of cataract surgery that might affect the results would include baseline comorbidities, previous surgeries, use of certain medications (i.e., alpha-adrenergic antagonist). Only relevant confounders should be incorporated in the models to prevent model overfitting. As already described, there should be at least 20 outcomes for every potential confounder included in the model. When identifying covariates, researchers should determine whether each covariate is highly correlated with other variables in the model, including the key predictor and the outcome variable, and include only 1 in a group of highly correlated variables.

A univariate analysis evaluates the relationship between the predictor and outcome variables without adjustment for confounders. Multivariable regressions come in multiple forms. Multivariable linear regression is used when the outcome of interest is continuous. Multivariable logistic regression is used when the outcome is binary (yes/no). A requirement for multivariable logistic regression is that all enrollees had continuous enrollment for a constant length of time. This is rarely the case with claims data because enrollees often enter and exit the plan at different times and thus have differing durations of observable time with risk for the outcome. To account for differing times in the plan among enrollees, researchers can perform Cox regression analysis, which assesses time to an event.

Disease Rate Comparisons with Population-Based Studies

Caution is warranted when comparing findings of disease incidence or prevalence in claims data with those generated from population-based studies. Potential reasons for differences in rates between these 2 sources include differences in disease definitions (reliance on billing codes vs. clinical criteria); study location (most population-based studies are conducted in specific communities, whereas claims data can contain a nationwide sample); level of expertise of those assessing the patients for disease; adjustment of confounders; and types of patients (those who enroll in population-based studies may generally be healthier than others are). Furthermore, incidence and prevalence rates often are captured in observational and population-based studies at one point in time (point-prevalence), whereas with claims data, patients are usually followed in the plan for years, generating period-prevalence estimates. These 2 types of estimates often are not directly comparable.[6,40,42]

Future Uses of Claims Data

The Affordable Care Act (Public Law 111–148) mandates that health care claims data be used extensively, beginning in 2015, to assess resource use and quality of care.[43] We anticipate that the reliance on claims data will increase, including its use to develop and implement alternative options for reimbursing providers (e.g., bundling of services for management of a given condition) and to judge quality and efficiency of care. This makes it incumbent on the eye care provider to use the most appropriate and accurate billing codes, especially in describing disease severity. Opportunities to link administrative claims data with clinical data may increase with the use of electronic health records, allowing for better assessment and quantification of quality of care and patient outcomes.

Convening a forum for health services researchers who analyze claims data may be useful—specifically, to reach consensus on the specific billing codes that should be used to identify patients with relatively common ophthalmologic conditions. Standardizing the codes used to research particular conditions would enable easier comparison of findings across various groups and data sources. Furthermore, researchers working with claims data should be encouraged to replicate their findings using other data sources (e.g., perform similar analyses involving Medicare patients and managed-care enrollees) to explore whether findings in one population are generalizable to another and to illuminate the extent to which differences in resource use or outcomes may be attributable to insurance type.

Despite limitations associated with claims-data analyses, these data sources will likely be used increasingly by insurers and other third-party payers to determine provider reimbursements and by policymakers to guide important decisions on resource allocation. Clinicians are encouraged to use newer billing codes to reflect disease severity and distinguish accurately among different conditions.

In conclusion, research involving claims data should be conducted with rigorous, sound methods, including those discussed in this article, and reported with sufficient methodological details. Readers are advised to look for these methodological details to appraise these reports carefully. We encourage researchers to join our conversation on ways to elevate the quality and rigor of ophthalmologic claims-data analyses.

Acknowledgments

References

1. Coleman AL, Morgenstern H. Use of insurance claims databases to evaluate the outcomes of ophthalmic surgery. Surv Ophthalmol. 1997; 42:271–8. [PubMed: 9406373]

2. Motheral B, Brooks J, Clark MA, et al. A checklist for retrospective database studiesdreport of the ISPOR Task Force on Retrospective Databases. Value Health. 2003; 6:90–7. [PubMed: 12641858]

3. International Classification of Diseases, Ninth Revision, Clinical Modification. American Medical Association; Chicago, IL:

4. Current Procedural Terminology 4, Professional Edition. American Medical Association; Chicago, IL:

5. Stein JD, Pasquale LR, Talwar N, et al. Geographic and climatic factors associated with exfoliation syndrome. Arch Ophthalmol. 2011; 129:1053–60. [PubMed: 21825188]

6. Stein JD, Kim DS, Niziol LM, et al. Differences in rates of glaucoma among Asian Americans and other racial groups, and among various Asian ethnic groups. Ophthalmology. 2011; 118:1031–7. [PubMed: 21310489]

7. Sloan FA, Picone G, Brown DS, Lee PP. Longitudinal analysis of the relationship between regular eye examinations and changes in visual and functional status. J Am Geriatr Soc. 2005; 53:1867–74. [PubMed: 16274366]

8. Coleman AL, Yu F, Evans SJ. Use of gonioscopy in Medicare beneficiaries before glaucoma surgery. J Glaucoma. 2006; 15:486–93. [PubMed: 17106360]

9. Stein JD, Hanrahan BW, Comer GM, Sloan FA. Diffusion of technologies for the care of olderadults with exudative age-related macular degeneration. Am J Ophthalmol. 2013; 155:688–96. [PubMed: 23219066]

10. Stein JD, Talwar N, Laverne AM, et al. Trends in use of ancillary glaucoma tests for patients with open-angle glaucoma from 2001 to 2009. Ophthalmology. 2012; 119:748–58. [PubMed: 22218146]

11. Stein JD, Ayyagari P, Sloan FA, Lee PP. Rates of glaucoma medication utilization among persons with primary open-angle glaucoma, 1992 to 2002. Ophthalmology. 2008; 115:1315–9. [PubMed: 18321581]

12. Ramulu PY, Corcoran KJ, Corcoran SL, Robin AL. Utilization of various glaucoma surgeries and procedures in Medicare beneficiaries from 1995 to 2004. Ophthalmology. 2007; 114:2265–70. [PubMed: 17466376]

13. Ramulu PY, Do DV, Corcoran KJ, et al. Use of retinal procedures in Medicare beneficiaries from 1997 to 2007. Arch Ophthalmol. 2010; 128:1335–40. [PubMed: 20938004]

14. Stein JD, Grossman DS, Mundy KM, et al. Severe adverse events after cataract surgery among Medicare beneficiaries. Ophthalmology. 2011; 118:1716–23. [PubMed: 21640382]

15. Stein JD, Ruiz D Jr, Belsky D, et al. Longitudinal rates of postoperative adverse outcomes after glaucoma surgery among Medicare beneficiaries 1994 to 2005. Ophthalmology. 2008; 115:1109–16. [PubMed: 18598818]

16. Stein JD, Zacks DN, Grossman D, et al. Adverse events after pars plana vitrectomy among Medicare beneficiaries. Arch Ophthalmol. 2009; 127:1656–63. [PubMed: 20008722]

17. Greenberg PB, Tseng VL, Wu WC, et al. Prevalence and predictors of ocular complications associated with cataract surgery in United States veterans. Ophthalmology. 2011; 118:507–14. [PubMed: 21035868]

18. Tseng VL, Yu F, Lum F, Coleman AL. Risk of fractures following cataract surgery in Medicare beneficiaries. JAMA. 2012; 308:493–501. [PubMed: 22851116]

19. Anastasopoulos E, Yu F, Coleman AL. Age-related macular degeneration is associated with an increased risk of hip fractures in the Medicare database. Am J Ophthalmol. 2006; 142:1081–3. [PubMed: 17157603]

20. Newman-Casey PA, Talwar N, Nan B, et al. The relationship between components of metabolic syndrome and open-angle glaucoma. Ophthalmology. 2011; 118:1318–26. [PubMed: 21481477]

21. Ostermann J, Sloan FA, Herndon L, Lee PP. Racial differences in glaucoma care: the longitudinal pattern of care. Arch Ophthalmol. 2005; 123:1693–8. [PubMed: 16344441]

22. Day S, Acquah K, Lee PP, et al. Medicare costs for neovascular age-related macular degeneration, 1994-2007. Am J Ophthalmol. 2011; 152:1014–20. [PubMed: 21843875]

23. Lam BL, Zheng DD, Davila EP, et al. Trends in glaucoma medication expenditure: Medical Expenditure Panel Survey 2001-2006. Arch Ophthalmol. 2011; 129:1345–50. [PubMed: 21670329]

24. Stein JD, Niziol LM, Musch DC, et al. Longitudinal trends in resource use in an incident cohort of open-angle glaucoma patients: resource use in open-angle glaucoma. Am J Ophthalmol. 2012; 154:452–9. [PubMed: 22789564]

25. Shea AM, Curtis LH, Hammill BG, et al. Resource use and costs associated with diabetic macular edema in elderly persons. Arch Ophthalmol. 2008; 126:1748–54. [PubMed: 19064859]

26. Coleman AL, Yu F. Eye-related Medicare costs for patients with age-related macular degeneration from 1995 to 1999. Ophthalmology. 2008; 115:18–25. [PubMed: 17572499]

27. Stoltzfus JC. Logistic regression: a brief primer. Acad Emerg Med. 2011; 18:1099–104. [PubMed: 21996075]

28. Patty L, Wu C, Torres M, et al. Validity of self-reported eye disease and treatment in a population-based study: the Los Angeles Latino Eye Study. Ophthalmology. 2012; 119:1725–30. [PubMed: 22537615]

29. Paikal D, Yu F, Holland GN, Coleman AL. Coding of glaucoma for patients with uveitis in the Medicare database. J Glaucoma. 2006; 15:13–6. [PubMed: 16378011]

30. Quigley HA, Friedman DS, Hahn SR. Evaluation of practice patterns for the care of open-angle glaucoma compared with claims data: the Glaucoma Adherence and Persistency Study. Ophthalmology. 2007; 114:1599–606. [PubMed: 17572498]

31. O'Malley KJ, Cook KF, Price MD, et al. Measuring diagnoses: ICD code accuracy. Health Serv Res. 2005; 40:1620–39. [PubMed: 16178999]

32. Lloyd SS, Rissing JP. Physician and coding errors in patient records. JAMA. 1985; 254:1330–6. [PubMed: 3927014]

33. Pawlson LG, Scholle SH, Powers A. Comparison of administrative-only versus administrative plus chart review data for reporting HEDIS hybrid measures. Am J Manag Care. 2007; 13:553–8. [PubMed: 17927459]

34. Fitch, K.; Bernstein, SJ.; Aguilar, MD., et al. The RAND/UCLA Appropriateness Method User's Manual. RAND; Santa Monica, CA: 2001. Available at: http://www.rand.org/content/dam/rand/pubs/monograph_reports/2011/MR1269.pdf. [January 19, 2013]

35. Dresser MV, Feingold L, Rosenkranz SL, Coltin KL. Clinical quality measurement. Comparing chart review and automated methodologies. Med Care. 1997; 35:539–52. [PubMed: 9191700]

36. Javitt JC, McBean AM, Sastry SS, DiPaolo F. Accuracy of coding in Medicare part B claims. Cataract as a case study. Arch Ophthalmol. 1993; 111:605–7. [PubMed: 8489437]

37. Bearelly S, Mruthyunjaya P, Tzeng JP, et al. Identification of patients with diabetic macular edema from claims data: a validation study. Arch Ophthalmol. 2008; 126:986–9. [PubMed: 18625948]

38. Latkany P, Duggal M, Goulet J, et al. The need for validation of large administrative databases: Veterans Health Administration ICD-9CM coding of exudative age-related macular degeneration and ranibizumab usage. J Ocul Biol Dis Infor. 2010; 3:30–4. [PubMed: 21139706]

39. Lee PP, Levin LA, Walt JG, et al. The impact of glaucoma coding in a large claims database. Am J Ophthalmol. 2007; 143:867–70. [PubMed: 17452173]

40. Sloan FA, Brown DS, Carlisle ES, et al. Estimates of incidence rates with longitudinal claims data. Arch Ophthalmol. 2003; 121:1462–8. [PubMed: 14557184]

41. Friedman DS, Quigley HA, Gelb L, et al. Using pharmacy claims data to study adherence to glaucoma medications: methodology and findings of the Glaucoma Adherence and Persistency Study (GAPS). Invest Ophthalmol Vis Sci. 2007; 48:5052–7. [PubMed: 17962457]

42. Stein JD, Vanderbeek BL, Talwar N, et al. Rates of nonexudative and exudative age-related macular degeneration among Asian American ethnic groups. Invest Ophthalmol Vis Sci. 2011; 52:6842–8. [PubMed: 21778274]

43. [November 4, 2013] H. R. 3590(111): Patient Protection and Affordable Care Act. Available at: http://thomas.loc.gov/cgi-bin/query/z?c111:H.R. 3590.pp:.

Table 1

Checklist for the Reporting of Studies That Use Claims Data[*]

Description of data source	Was a description of the sociodemographic characteristics and health care profile of the population provided? Did the investigators acknowledge limitations of services provided due to type of insurance or plan type of the enrollees, benefit design, and how providers are reimbursed for services studied?
Checks of data quality	Sources of unreliable data include changes in reporting/coding practices over time and in reporting resulting from changes in reimbursement, and services may be inadequately captured if not covered by the plan. Did the investigators explain how they handled missing and out-of-range values? Did the investigators explain how they handled duplicate claims and inconsistencies (differences in age of same patient on different claims)? Did the investigators compare the reported rates of disease or use with established norms or other data sources? If other researchers have studied the reliability and validity of the data source used, those should be cited. Have the necessary linkages among data sources or sites of care been carried out appropriately? Is there an explanation of how member eligibility was determined?
Sample selection	Is there a sample selection figure to easily show readers the numbers of enrollees included and excluded and for what reasons? Is there is justification provided for using the chosen inclusion/exclusion criteria for selecting beneficiaries for the study sample? Is there a transparent listing of all of the ICD-9-CM and CPT codes used in the study? Were enrollees who were noncontinuously enrolled in the health plan during the entire study period included in the analysis?
Analysis	Is the data analysis plan clearly described? Were research hypotheses generated a priori or were the findings generated the result of unsystematic data exploration? Did the investigator provide a cogent rationale for the study design chosen, in light of the data, setting, and research questions? Are limitations of the study design chosen clearly delineated to the reader? Examples of potential biases include selection bias, maturation, and regression to the mean. For studies reporting treatment effects, was there a control group created to compare against the group receiving the intervention? Did the investigators censor subjects, and, if so, did they explain how this may affect the sample selection or generalizability of the cohort? Are the end points or outcomes clearly defined on the basis of diagnosis or procedure codes or other criteria? Did the investigators justify the definition of the end points they chose for the analysis or cite other sources who used similar criteria? Were sensitivity analyses performed to explore the impact of changing the criteria for study inclusion or the definition of the outcome(s) of interest? Is there a temporal relationship between the exposure and the outcome of interest (did the researchers require the exposure to come before the outcome)?
Statistics	Were important confounding factors identified and adjusted for in the analyses either by stratification of the sample by the confounding variable or by the use of multivariable statistical techniques? What sort of risk adjustment was performed? Did the investigators account for differences in sociodemographic characteristics, medical comorbidities, disease severity? Were adequate tests of the statistical assumptions performed? Examples include testing for multicollinearity and adjustment for multiple comparisons.
Discussion/conclusions	Did the investigators provide a rationale for the study findings in light of the existing literature? Were alternative explanations for the findings offered? Did the investigators comment on the clinical or economic relevance of the study findings because statistical significance may not necessarily translate into clinical significance? Did the authors address concerns about the generalizability of study findings to other groups?
Funding sources	Were the funding sources for the analyses clearly identified? Did the funding sources participate in designing or conducting the study?

CPT = Current Procedural Terminology; ICD-9-CM = International Classification of Diseases, 9th Revision, Clinical Modification.

[*] Adapted from Motheral B, Brooks J, Clark MA, et al. A checklist for retrospective database studies—report of the ISPOR Task Force on Retrospective Databases. Value Health 2003;6:90-7.[2]