The Use of Claims Data in Healthcare Research

Kari Ferver, Bryan Burton and Paul Jesilow[*]

Department of Criminology, Law and Society, School of Social Ecology, University of California, Irvine, CA 92697-7080, USA

Abstract: *Background*: Claims-based studies have become common during the past 15 years. The electronic records of millions of transactions include information entered on bills (claims) submitted by healthcare providers to third-party payers. They are an attractive data source to researchers; however, they contain many weaknesses that threaten the integrity of many studies that utilize them.

Purpose: We wanted to learn: (1) how widespread are studies that utilize claims data; (2) in what healthcare areas are claims data being used; (3) is the use of claims data increasing; and (4) do researchers inform readers of the weaknesses of the data?

Data Sources: 1,956 original research studies, published during 2000-2005 in five healthcare journals, were reviewed to: (1) determine their data sources, (2) establish their healthcare areas, and (3) if claims-based, to ascertain any discussion of the data's weaknesses.

Conclusions: Use of claims databases in research may have leveled. They are often used to study healthcare areas for which they may be appropriate, but they are also used in areas where they might not be suitable. Less than half the authors of claims-based research mentioned any weaknesses of the data.

INTRODUCTION

Hundreds of medical research studies published during the past 25 years have relied on claims data. Claims data, also called billing or sometimes administrative data, have been used to study such diverse issues as the use of comorbidity indices to predict risk of death [1], antiretroviral therapy [2], psychotropic drug usage [3], children's mental health services [4], substance use disorders [5], the cost effectiveness of lung-volume–reduction surgery [6], diabetes preventive services [7], and numerous other areas [8-11]. Most recently, they have been used as data by health and life insurance companies to reach decisions about whether to cover specific individuals [12].

Claims databases, simply put, are electronic records of millions of transactions that are purported to have occurred between patients and healthcare providers. They include information entered on bills (claims) submitted by hospitals, clinics, nursing homes, pharmacies, individual providers, and other medical professionals to public (e.g. Medicare and Medicaid) and private (e.g. Blue Cross/Blue Shield) insurance entities. The frequency and growth of such studies may impact healthcare policies that rely on them. Claims data may provide useful information, but they are not perfect and the extent of weaknesses inherent in them brings into question the validity of some studies that utilize them. The research reported here was designed to help us understand whether concern about the use of claims data in healthcare

studies is ill placed or whether such concern is justified, or, as with many things, that the answer lies somewhere in-between the two poles.

The Advantages of Claims Data

Claims data are appealing to researchers because they offer numerous advantages. They are anonymous, plentiful, inexpensive, and widely available in electronic format [13]. This has made their use common in research studies, particularly as a replacement for medical records.

The anonymity of claims data is a major advantage in some studies. There use, for example, may be beneficial in research that is attempting to learn sensitive information about groups. Such information could be obtained directly from patients by surveying or interviewing them, but studies suggest that individuals may not accurately self-report; for example, only 2% of syphilis sufferers reported the condition in an early study [14]. Claims data also do not require patient authorization for use and are free from non-response and dropout, which are problems for studies that rely on surveys or interviews for their data [15].

Claims databases are also particularly useful for finding sizable groups of patients with rare conditions, such as quadriplegia or aplastic anemia, who might be difficult to locate by other means [16]. Claims data may also be good for establishing the cost for certain diagnoses. The purpose of a claim is to collect payment, so it is convenient for researchers to consult fee schedules and reimbursement data and perform cost-effectiveness analyses [e.g., see 17]. Coding errors and questionable billing practices may contaminate the costs of a specific diagnosis, but the financial outlays are nevertheless real.

*Address correspondence to this author at the Department of Criminology, Law and Society, School of Social Ecology, University of California, Irvine, CA 92697-7080, USA; Fax: 949 824 3001; Tel: 949 856 0677; E-mail: paul.jesilow@uci.edu

A major use of claims data is as a substitute for the information contained in patients' medical records. Medical records are often not available in electronic format, which makes them harder to access. Researchers must hire professionals and train them to obtain from the medical records the bits of information (such as age or presence of a disease) that are pertinent to the research [18]. This process, called abstracting, requires skilled individuals to pay close attention to details and expertly interpret information in medical records. Such individuals, often doctors or advanced medical students, are expensive, especially in studies that have multiple individuals abstract each file to ensure accuracy. The result is that research that utilizes medical records must minimize the number of sites that are studied, which may weaken the validity and reliability of the findings. Claims-based studies can be much more inclusive at a greatly reduced cost.

Weaknesses of Claims Data

Problems with claims data stem from the fact that they were fashioned to obtain reimbursement and were not designed for the purposes of researchers. Institutional providers, such as hospitals, bill using the diagnosis-related group (DRG) prospective payment system, which classifies patients into one of hundreds of groups based upon their principal diagnoses and complications. Reimbursement is based on the International Classification of Diseases, Clinical Modification (ICD-9-CM) coding system and is fixed at a certain rate; it is assumed that individuals with the same diagnoses on average will require similar procedures and treatments [19].

Information is included on claims forms primarily because it is necessary for compensation and non-essential data for billing is usually excluded. This fact impacts the usefulness of claims data for researchers. Consider the recent announcement from the Centers for Medicare and Medicaid Services (CMS) that it would not pay for treating some preventable conditions, including transfusions with the wrong blood type, blood clots in the leg following knee or hip-replacement and complications stemming from poor control of blood sugar levels. CMS has tentative plans to extend its do-not-pay list (commonly referred to as "never events") to additional conditions that it believes are the result of medical errors. The goal of the agency is to reduce such events by providing a fiscal disincentive for their occurrences [20]. The lack of reimbursement for these preventable conditions, however, will likely result in diminished reporting of them in claims, and make the data problematic as a measure for errors. CMS has added some quality measures to the existing list that hospitals are mandated to report to maintain full reimbursement levels, and this may, in part, offset the loss of information in the claims data [20]. But their addition is unlikely to fully correct the problem; billing clerks will likely not include conditions for which the hospital cannot be reimbursed (and which may reflect poorly on the quality of their employer).

Some information are missing from claims because the medical procedures are commonly bundled into claims for office visits [21], or medical providers may not bother to bill some procedures because the reimbursements are perceived to not be worth the effort [22]. Recognition that bundling is a problem in some claims-based studies has limited their use. Pediatric immunizations and prenatal care visits are illustrative; they have often been bundled with regular office visits for billing purposes, causing them to be vastly underreported in claims data [21].

Comparatively minor tests or routine hospital procedures that are normally associated with specific diagnoses are less likely to be recorded on a claims form because they are unlikely to receive additional reimbursement if billed separately [23]. Chronic diseases are often underreported in claims for outpatient visits because often only one diagnosis will suffice for reimbursement. Underlying causes for patients' current visit are unlikely to be recorded [24]. Information on medications is regularly absent from hospital claims [25]. Non-operating room activities are often not noted because guidelines only require recording procedures that are surgical in nature, require specialized training, or carry a procedural or anesthetic risk [26]. Conditions and procedures that do not need to be listed for reimbursement purposes are omitted [27]. Comorbidities, for example, are underreported and this may lead to inaccurate research results. Mistaken expected death rates in one study were in part attributed to underreporting of comorbidities in the hospital data [28].

Part of the problem for underreporting on claims is the result of the form that has been used for billing. It restricted information by providing space for a limited number of procedures and diagnoses. The form has been expanded in response to this recognized limitation, but this does not address the underlying weakness; there is a financial disincentive for the additional information to be recorded. Extra time spent by billing clerks on the forms will be lost income to the institution. It does not receive additional compensation as a result of the labor. It should not be surprising that hospital personnel, who are taught to expeditiously complete the forms, are unlikely to risk the wrath of supervisors in order to satisfy external desires for additional information.

Clerks introduce additional drawbacks while abstracting the medical records to produce the bills/claims. Coders commonly enter incorrect information, as the result of faulty decisions about what to code, misreading of the medical record, and typographical errors [29, 30]. The more ambiguous the decision (gray areas), the more likely the coder is to provide inaccurate information [31-33]. This result is not always by accident, but the outcome of systemic efforts.

Claims exist for the purpose of generating income for individual and institutional providers and those who produce them are primarily guided by this function. Billing staff receive special training in classes taught in vocational schools or online classes. They learn how to interpret information in medical records to determine diagnoses or procedures and to match this information to a DRG and translate it into ICD-9-CM codes. They receive instruction on how to handle ambiguous information in the medical record and this leads to inaccuracies in claims data. They learn to bill for what is "reasonable," which their instructors teach is the code that will provide the highest reimbursement that can plausibly be supported by the medical record.

Consultants who are hired by physicians and healthcare institutions to advise them on government reimbursement add to inaccuracies in the claims data. A United States General Accounting Office report on the subject noted that when it came to billing, the consultants' "emphasis was not that the code selection be correct or even that the services be performed, but rather that it is important to create a documentary basis for the codes billed in the event of an audit. . . the documentation created is the support for billing for services at higher code levels than warranted [34]". Moreover, consultants encouraged medically unnecessary tests and procedures to create documentation for higher levels of reimbursement [34]. Such efforts to increase income weaken the usefulness of claims data as measures of healthcare and the problem is extensive. A government report, based on a sample of Medicare recipients, estimated about $2 billion per year in overpayments to FFS Medicare providers as a result of coding errors and an additional $5 to $8 billion as the result of medically unnecessary services [35].

Fraud may be a common cause for inaccuracies in claims data [36]. There is substantial evidence that illegal billing practices occur in the healthcare industry, and that these practices affect billing data [37-40]. Illegal billing practices range from the idiosyncratic – kickbacks, billing for services not provided, ordering unnecessary tests and treatments [41] – to more industry-wide behaviors, such as upcoding [42]. Upcoding, in particular, can negatively impact claims data.

Upcoding undermines the usefulness of claims data as measures of healthcare because patients appear to be sicker in the claims data than they actually were [43]. Upcoding in hospitals (also known as "code creep" or "DRG creep") is common [38, 44-50]. The practice involves charging for a more expensive DRG than the legitimate one. Bruce Psaty and his colleagues [48] reviewed the charts of patients with a Medicare discharge diagnosis of heart failure. The researchers found that in more than one-third of the cases they reviewed, there was no evidence to support the hospitals billing claims that the cases were worthy of a higher fee. In another study of cerebrovascular disorders, hospitals typically billed for the incorrect, higher-reimbursed DRG [51].

There are other illegal industry-wide practices that have undermined the legitimacy of claims data for use in policy studies [52]. Patients, in one scheme, are moved from hospitals to outpatient or other settings in order to maximize payments from the insurer. The patients' movements are guided by profit and not by medical decisions. Information submitted in claims to the insurers in order to justify payments, however, will fraudulently reflect medical reasons for the moves. The effect of this and other illegal schemes is to undermine the legitimacy of the claims data as an accurate reflection of the health of the patients or the care that was delivered to them.

The weaknesses of claims data have led a number of researchers to scrutinize the concordance between them and medical records (as well as registries) for a number of healthcare issues, including mental health services [53], comorbidities [26, 54, 55], antidepressant use [56], breast cancer-related surgeries [57], surgical and medical complications [58, 59], cardiovascular and stroke risk factors

[60], heart failure [47], cancer surveillance [61], pneumonia [62], and high-risk conditions for influenza [63]. The highest concordance between the medical record and claims data have been recorded for major procedures, such as mastectomies [26, 31, 32, 57]. The concordance rate is worse for conditions that may result in conflicting diagnoses, such as mental health issues [64] or for minor procedures [26].

Results obtained from some studies that utilize claims data may be of debatable value and their use raises for us four questions that we feel are important to answer. First, how widespread are claims-based studies? It seems reasonable to learn the parameters of a potential difficulty before deciding whether substantial effort is needed (in the vernacular, there is no reason to make a mountain out of a molehill). Second, is their use increasing? This question is similar in purpose to our first one. Third, in what healthcare areas are claims data being used? We wanted to gain a general portrait of their use. Finally, do researchers inform readers of the weaknesses of the data? The authors of past claims-based studies may have understood some of the weaknesses of the data, but failed to mention them. Policymakers might rely on the studies without considering the weaknesses that might undermine the validity of the results.[1] We wanted to learn if non-reporting was common and worthy of further study.

METHODS

Reviewing published healthcare studies is the best method to obtain data to answer our questions, but we could not possibly review all healthcare journals published since the introduction of claims data in order to resolve our questions. The task would be overly time-consuming and expensive given the value of the results. There are, for example, more than 450 full-texts online journals that include the word "health" in their titles.

A common strategy for conducting reviews is to examine a sample of articles obtained *via* an academic database, such as PubMed. We initially tried this approach, but it quickly proved ineffective for answering our questions for a number of reasons. First, it was impossible for us to determine the number of published studies that used claims data in any single recent year. Studies do not uniformly use "claims data" as a keyword and the words do not always appear in the text of an article. Sometimes the data are referred to as "administrative data" or as data derived from Medicare or Medicaid reports or from private insurance or the name of an insurer or from pharmaceutical data, as well as other terms. Even "hits" for "claims data" did not guarantee that the study utilized claims data. After initially reviewing scores of articles (only a portion of which as it turned out used claims data) it became apparent that reading the thousands of articles produced by our searches would not help us understand the extent of claims' studies or whether their use was increasing. We determined to take a more systematic

---

1 Areas of study, other than healthcare, have had similar problems. Throughout much of the 1960s and 1970s, criminological research on the impact of punishment on the extent of crime relied on official government statistics, without mentioning the weaknesses of the data. Policymakers used the studies to support their positions until a National Academy of Science publication discounted use of the government data [65].

approach that would better lend itself to answering our questions.

We decided to focus our attention on a selected sample of journals. We could then review all the journals' published studies during a specific time frame. This would allow us to calculate a rate for studies that utilized claims data, which could then be used for comparisons to answer our questions.

To begin we reviewed scores of journals to determine if they had any relevant articles, a necessity since we hoped to determine if there had been some variation in the percentage of studies utilizing claims data that are published year-to-year. We wanted the journals to also vary with respect to the extent that they published studies that utilized claims data. Selecting only journals that rarely published claims studies would result in underestimating the data's use. Picking journals that often published them would produce the opposite. We finally chose five health/medical journals for study: *The New England Journal of Medicine,* the *American Journal of Medical Quality, Medical Care, Medical Care Research and Review*, and *Health Care Financing Review*. These five journals represent a continuum with respect to the use of claims data in research studies.

The New England Journal of Medicine was selected to represent the first tier of clinical research journals. The initial review revealed that the journal published relatively few studies that utilized claims data, but there were some. Excluding the title would result in an overestimation of the use of claims data in research. The journal is owned and published by the Massachusetts Medical Society. It is highly regarded and frequently cited; it was the eighth most cited science journal in 2006 [66] and its impact factor that year was 51.296. It publishes on a variety of topics of importance to biomedical science and clinical practice, with an emphasis on internal medicine and specialty areas [67]. These topics, with their cutting-edge focus on diseases and treatments, are well-suited for clinical trials, but not for evaluative studies with claims data, especially for newer treatments that may not yet be covered by health insurance and as a result may be missing from claims data.

The *American Journal of Medical Quality* is the official publication of the American College of Medical Quality, which offers, according to its website, "national and international resources for providers, payers, purchasers and legislators [68]". The impact factor of the journal in 2006 was 1.046. The goal of the journal is to keep "readers informed of the resources, processes, and perspectives contributing to health care services." Topics revolve around measuring the quality of healthcare through patient satisfaction measures or evaluations of healthcare delivery and management [69]. The availability of claims data may be attractive to researchers studying widespread management and delivery trends; other data sources may be limited by geography or other factors. Our initial review of the journal found numerous articles utilizing claims data.

Medical Care is the official journal of the medical care section of the American Public Health Association, an organization devoted to protecting communities from preventable diseases and promoting preventive care. The journal is devoted to all aspects of the administration and delivery of healthcare, and its published studies focus on

"issues related to the research, planning, organization, financing, provision, and evaluation of health services" [70]. The impact factor of the journal is 3.745; it ranks third of 56 in the Health Care Sciences and Services category, and tenth of 98 in the Public, Environmental & Occupational Health category.

Medical Care Research and Review describes its goal as providing essential information about the field of health services to researchers, policy makers, managers, and practitioners, with focus on topics in "organization, financing, health care reform, quality of care, and patient-provider relationships" [71]. The journal's 2006 impact factor was 2.091. The two journals, *Medical Care* and *Medical Care Research and Review* have broad, overlapping areas of interest, and we chose them because we expected studies published in them to use a wide range of data, including claims.

Health Care Financing Review is published by the Centers for Medicare and Medicaid Services, a United States' governmental organization that provides healthcare coverage through the Medicare and Medicaid programs. The journal's goal is to "contribute to an improved understanding of the Medicare and Medicaid Programs and the U.S. health care system by presenting information and analyses on a broad range of health care financing and delivery issues" [72]. The journal's 2006 impact factor was 1.027. It was selected because our initial review revealed it published substantial numbers of claims-based studies.

Types of Studies

Claims-based quantitative studies were but one category of investigations published in our selected journals. They also relied on data derived from a variety of other methodologies, including data derived from clinical trials, medical record chart abstractions, surveys and interviews, and a category we call "other administrative data".[2] We initially categorized the studies by these data categories. We wanted to determine each journal's trend with respect to this matter. Each data source has strengths and weaknesses. Any rise in the percentage of studies utilizing claims data must be considered in light of concomitant declines in studies that utilize other data. This is an important aspect of the current work. A decline in studies that utilized data from clinical trials in favor of ones that employed claims data, for example, would not be positive.

Clinical trials are reliable for medical decision-making and are considered the overall gold standard in medical research [73]. Clinical studies are experimentally grounded and are meant to increase the validity of the results by isolating the effects of a treatment or other intervention. Clinical trials do have limitations. Trials do not reveal long-term effects, or any impacts on sizable portions of the population that were not included in the clinical trial, or outcomes that might be revealed in actual practice; clinical trials are often designed to determine whether a treatment does better than a placebo under controlled circumstances [33, 74, 75].

---

2 The journals also published editorials, commentaries, and literature reviews. These were excluded from our analyses.

Medical records consist of "what is heard, seen, thought and done concerning the patient" [76]. They provide rich, detailed clinical information on diagnoses, procedures, risk factors, and outcomes of care in practice [77], but they have their limitations. They are expensive and difficult to obtain due to privacy issues. In addition, there may be accuracy problems caused by incorrect documentation, illegible notes, missing lab reports, and bad chart abstractor skills [13]. Medical records can have errors that are introduced by physicians, including failures to list procedures or diagnoses, inappropriate primary diagnoses, and inadequate or vague terminology [78-80]. Physicians are attuned to the information needed for billing, and will at times not include in medical records information that is not relevant for reimbursement. One study found that physicians failed to record the things that they did in patients' charts about half the time and the authors of the study, as a result of their findings, discouraged the utilization of chart abstraction for quality of care studies [81]. Physicians may put false information in the medical record [82-86]. One survey, for example, found that about half of the physician-participants were willing to deceive third-party payers on matters associated with coronary bypass surgery, arterial revascularization, and intravenous pain medication and nutrition so that their treatments of the patients would be reimbursed that would otherwise have been denied [87]. Still, it is reasonable to assume that medical records are better indicators of patients and the healthcare they received than are claims data, which are yet another step removed from the institution/physician-patient interaction; if medical records are themselves inaccurate, any abstraction of them, including their use for the production of claims, can only be erroneous.

A relatively new source of data come from a category we have entitled "other administrative data". They include hospital discharge abstracts, registries or research networks, such as cancer registries or the HIV Research Network [e.g., see 88] Veterans Affairs' data [e.g., see 89] and combinations of other measures [e.g., see 90]. These data frequently have problems that are similar to claims data. Comorbidities, for example, are often underreported in hospital discharge abstracts [91], abstractors are not consistent from one site to another [92-94] and there are differences between these abstractions and that which is in the medical record [95].

For us, a major difficulty was determining if the administrative data base(s) relied, even in part, on claims data. At times we were able to conclude that at least one measure of a composite was based on claims [e.g., see 25] and we coded such studies as claims, since those data are of primary interest to us. We often, however, were confused about the source of information in a data set. Hospital discharge abstracts are illustrative. California's Patient Discharge Data Set, for example: includes an abstract of every discharge from every non-federal, licensed hospital in California. Each abstract lists the patient's birth date, sex, race, ZIP code, encrypted social security number, source and type of admission, discharge disposition, expected principal source of payment, total charges, principal diagnosis and up to 24 secondary diagnoses, up to 21 procedures, and up to 5 external causes of injury [94: pg 718].

Much of the information included in such hospital discharge abstracts should be from the medical record [e.g., see 94]. For us, however, the question must be from where did matters such as "expected principal source of payment" and "total charges" develop. Further complicating matters for us, these cost issues are irrelevant in many studies that utilize discharge abstracts, since many deal with quality of care issues. We decided to categorize such studies as other abstractions of the medical record unless it was clear that claims (billings) played some role in the research.

Surveys and interviews are another source of data in healthcare studies. They are commonly used as methods for gathering subjective opinions, rather than objective facts [96]. Mail and telephone surveys are capable of collecting data that are not available by other methods, such as measures of patient satisfaction that are unlikely to be collected in a clinical trial or appear in medical records or on claims forms. An increase in the use of surveys or interviews may indicate a trend towards studying different aspects of healthcare than those that might be revealed from other data sources. General problems with surveys include low response rates, poor recall by respondents, failure by respondents to provide information on topics they find too sensitive, and attempts by respondents to provide answers that they believe the researchers want [96-99].

Interviews are similar to surveys in that they are able to collect information and opinions directly from the patients that are not available from secondary data sources. They almost always involve non-experimental designs. Interviews have high response rates and allow the interviewer to collect supplementary information in order to clarify answers. But bias is very difficult to avoid in interviews because even well-trained interviewers provide both verbal and nonverbal cues that may influence what participants say. Participants, too, may try to please the interviewer and provide answers that they believe are desired. Interviews also have anonymity issues; healthcare research can be especially problematic because respondents may feel too exposed and intimidated to provide sensitive medical information [96, 100]. There is evidence, for example, that claims data do a better job of identifying patients who have had a specific procedure than do surveys or interviews [e.g., see 101]. Finally, interviews are expensive because personnel must be trained and supervised, interviews may need to be transcribed, and traveling to interviews consumes both time and money.

Some studies published in the five journals we reviewed employed more than one type of data. Several studies used claims data to find a study population and then interviewed or surveyed that population [e.g., see 102]. The use of claims data to find a study population avoids internal validity problems associated with other sampling methods (e.g., "snowball" samples), and since the claims data were not used as healthcare measures the accuracy problems associated with them are diminished. Other studies utilized more than one type of data as cross-checks, such as studies that compared the concordance of patients' self-report data with data obtained from medical records or claims [55, 56]. Multiple methods can greatly enhance the reliability of the data [103].

We carefully categorized studies that utilized multiple types of data. We prioritized the use of claims for coding.

Studies, for example, that used claims to identify beneficiaries and then used medical records for disease identification and analysis were coded as employing claims data. Interviews and surveys, because they are similar in ways described above, were collapsed into one category for analytical purposes. A small number of studies utilized medical records combined with either interviews or surveys. We coded these studies as employing medical records data. Medical record data, from our perspective, might be more useful for informing evidence-based healthcare and we prioritized its use for coding purposes.

Area of Study

We created healthcare categories in which to sort the different studies in order to determine if specific data sources were associated with certain topics. We categorized the studies as: (1) access to healthcare, (2) prevention and detection of disease, (3) quality of healthcare, (4) morbidity and mortality, and (5) interventions, therapies, and treatments. These are broad categories and may have overlapping characteristics. They are meant to create a general picture of the studies that utilized claims data.

The category "access to healthcare" encompasses studies that analyzed the availability of, or impediments to, obtaining healthcare, including factors such as race [104], continuity of care [105], and economic factors [106]. Claims data may be useful for some access-based studies. Race, for example, is coded more accurately in claims than are diagnoses or procedure codes [16]. But other data, for example self-reports, have been shown to be more accurate with respect to such basic demographic information.[3]

The category "prevention and detection of disease" includes studies that focus on preventive care and efforts to detect diseases through screening procedures. Claims data may not be appropriate for some of these types of studies, as procedures are often not accurately coded. Medicaid claims data, because of bundling, are poor indicators of whether children have been fully immunized [107]. Many inpatient screening procedures are not coded accurately in claims data; MRIs, for example, are rarely coded reliably [23]. This poor performance, however, is not universal; one study found that Pap smear claims data have more than 95% concordance with the medical record [108].

"Quality of healthcare" was broadly defined to include matters associated with physician accuracy in diagnosis and treatment (e.g. did the diagnosis correspond with the symptoms; did the diagnosis and treatment match) as well as the quality of patient healthcare (e.g., do patients, given their conditions, receive all recommended treatments and no unnecessary or harmful ones) or patients' satisfaction with their healthcare. We also included studies that tested the accuracy of different measures of healthcare quality. Claims data, with their many accuracy issues, may not be suitable in all circumstances for quality evaluations.

The category "morbidity and mortality" included studies that examined the rate of illness or death that arose from medical conditions. This category has some similarity to the quality of healthcare category, in that illness and death rates are sometimes used as quality of care measures, especially for hospital performance. Diagnosis, a necessary element for establishing morbidity or mortality, can be inaccurate in claims data [15, 24] and, as a result, claims data are not uniformly recommended for these types of studies. Diagnoses in medical records are more accurate and contain more information on patients' illnesses than their corresponding claims, which tend to underestimate certain disease rates (e.g., minor conditions that need not be listed for compensation) while likely overestimating others due to upcoding [46].

Our category--interventions, therapies and treatments-- covers a wide variety of activities that are aimed at improving specific health problems. Utilization of claims data are problematic for this type of study and may require substantial efforts to make them usable [13, 30]. Physicians may choose to not bill for everything they do or may take actions that result in treatments not appearing in claims (e.g., by giving patients drug samples instead of prescriptions). Hospital billing clerks may muddy data by choosing to exaggerate treatments in order to increase revenues. Medical records have shown to be more accurate than claims data in most studies that would fall in this category [22, 56, and 111].[4]

Data Collection

We culled information from 1,956 original research studies published between the years 2000-2005 in the New England Journal of Medicine, The American Journal of Medical Quality, Medical Care, Medical Care Research and Review, and Health Care Financing Review. The six years, 2000-2005, were selected to represent the current situation.[5] We recorded some basic information for each study (the title, authors, journal, volume number, year, and page numbers), as well as written descriptions (and codes) for the data type, medical area(s),[6] and whether the authors of studies that utilized claims acknowledged any limitations of the data. Studies that used claims and noted at least one limitation/weakness of the data were further reviewed to determine which limitations were mentioned.

The distribution of the 1,956 studies among the journals is not equal. *The New England Journal of Medicine* is published weekly and as a result 859 studies (43.9% of the total) appeared in this title. *Medical Care* is available monthly and occasionally publishes supplemental issues. We reviewed 718 studies (36.7% of the total) that appeared within its pages. *Health Care Financing Review* is a

---

[3] One study found only 60% agreement between claims and patient self-reports for race and ethnicity. The authors attributed most of the discordance to race being marked as 'unknown' in some of the claims data [109]. Another study found that survey and claims concordance rates varied based on the patient's race; white patients were classified correctly in the claims data 77% of the time, while other races were classified correctly between 1.4% and 76.4% of the time [110].

[4] Albeit the best methodology for interventions, therapies, and treatments, at least initially, is a randomized controlled trial. Clinical trials are explicitly designed to test the effectiveness of a treatment, and the internal validity of a trial's results cannot be matched by medical records or claims data.

[5] We knew, from a keyword search for "claims data" on the Web of Science that the use of claims data had increased throughout the 1990s (although such a general search cannot provide needed specifics).

[6] If a study fit into multiple categories, all applicable categories were initially considered; a senior researcher later recoded the study to reflect the dominant medical area.

quarterly publication. We reviewed 157 studies (8% of the total) that it published. The bimonthly *American Journal of Medical Quality* also occasionally provides supplements. We examined 127 studies (6.5% of the total) that appeared in this journal. Finally, *Medical Care Research and Review* comes out bimonthly, with occasional supplements. It provided 95 studies (4.8 % of the total) for review.

Limitations

Our study has several limitations which we tried to minimize. We looked at only five journals published during a six-year period. Journals were selected, however, to provide a general representation of journals that use claims data. There may be some coding errors, but the data went through three levels of review to minimize idiosyncratic mistakes. No inter-rater reliability measures were collected. Rather, we used multiple levels of review, the goal of which was to achieve uniformity and accuracy. Disagreements or questions about coding at the first two levels were submitted to the senior researcher, who discussed the matter with the other authors before reaching a decision.

RESULTS

The results reported here are designed to answer four questions: (1) how widespread are studies that utilize claims data; (2) in what healthcare areas are claims data being used; (3) is the use of claims data increasing; and (4) do researchers inform readers of the weaknesses of the data?

How Widespread are Studies that Utilize Claims Data?

We expected that there would be wide variation between the five journals concerning the percentage of studies in each that used claims data and this was the case. In general, the distribution followed the anticipated pattern. *The New*

England Journal of Medicine published by far the greatest percentage of studies derived from clinical trials, 75%; only 1.2% of studies published in the journal utilized claims data. Studies in *Health Care Financing Review*, at the other end of the continuum, utilized clinical trials less than 2% of the time, while 22.9% of the studies utilized claims data. There was, as expected, a statistically significant variation between the types of data found in each journal (Phi = .698 p <.001; see Table **1**). The uneven distribution in the five journals of the 1,956 studies results in a plurality of studies (38.9%) obtaining their data from clinical trials; 30.7% utilized survey and/or interview data; 18.5% involved data from medical records; 7.3% employ claims data; 4.6% utilized other administrative data (see Table **1**). It is important to note that a different set of journals would reveal a different distribution.

We removed *The New England Journal of Medicine* from only this initial analysis in order to focus on the journals that published studies that utilized more varied sources of data than it did. The statistically significant difference between the remaining four journals concerning the source of data for their published studies did not disappear, although there was consistency on some matters. Surveys and interviews were the favorite source of data at each of the remaining journals (45% of the studies utilized these data) and other administrative data remained the source of data for the smallest proportion of studies (7.75%) once articles published in *The New England Journal of Medicine* were removed from our analysis (clinical trials supplied the data for 10.6%; claims data accounted for 12.1%, and medical records were the data of choice in 24.5%).

In what Healthcare Areas are Claims Data being Used?

The type of data utilized in studies varied by the medical area being explored (Phi = .617, p< .001; see Table **2** for the

Table 1.   Journal by Type of Data

		Type of Data					Total
		Other Abstraction	**Clinical Trials**	**Surveys and Interviews**	**Medical Records**	**Claims**	
Medical Care	Count	55	91	305	194	73	718
	% within Journal	7.7%	12.7%	42.5%	27.0%	10.2%	100.0%
Health Care Financing Review	Count	16	3	82	20	36	157
	% within Journal	10.2%	1.9%	52.2%	12.7%	22.9%	100.0%
Medical Care Research and Review	Count	5	2	57	14	17	95
	% within Journal	5.3%	2.1%	60.0%	14.7%	17.9%	100.0%
New England Journal of Medicine	Count	4	645	107	93	10	859
	% within Journal	.5%	75.1%	12.5%	10.8%	1.2%	100.0%
American Journal of Medical Quality	Count	9	20	50	41	7	127
	% within Journal	7.1%	15.7%	39.4%	32.3%	5.5%	100.0%
Total	Count	89	761	601	362	143	1956
	% within Journals	4.6%	38.9%	30.7%	18.5%	7.3%	100.0%

Phi = .698 p < .001.

Table 2.   Type of Data by Medical Area

		Medical Area					Total
		Access	**Prevention**	**Quality**	**Morbidity**	**Treatment**	
Other Abstraction	Count	35	1	32	15	6	89
	% within	39.3%	1.1%	36.0%	16.9%	6.7%	100.0%
Clinical Trials	Count	24	101	69	183	384	761
	% within	3.2%	13.3%	9.1%	24.0%	50.5%	100.0%
Surveys & Interviews	Count	191	49	257	58	46	601
	% within	31.8%	8.2%	42.8%	9.7%	7.7%	100.0%
Medical Records	Count	98	33	112	70	49	362
	% within	27.1%	9.1%	30.9%	19.3%	13.5%	100.0%
Claims	Count	70	8	34	13	18	143
	% within	49.0%	5.6%	23.8%	9.1%	12.6%	100.0%
Total	Count	418	192	504	339	503	1956
	% within	21.4%	9.8%	25.8%	17.3%	25.7%	100.0%

Phi = .617 p < .001.

distribution of the type of data by medical area). Claims data were primarily used in studies that focused on access to healthcare (49% of claims-based studies) followed by quality issues (23.8%). Examples of access to healthcare studies included the effects of different societal or policy changes on patient use of health services. For example, one study used claims data to examine the effect of a reduction of welfare payments on mental health service use [112], while another study explored Medicare beneficiaries' usage of prescription drug discount cards [113]. Examples of quality of care studies included the use of claims data to establish benchmark left ventricular ejection fraction test rates for Medicare recipients [114], while another assessed the association between office systems and variations in diabetes care [7].

Claims data were less likely to be used to study morbidity issues (9.1% of claims-based studies), such as patterns of diagnoses, re-hospitalization rates, survival rates, and other large-scale measures associated with morbidity and mortality [115]. Similarly, the data were less likely to be used to examine matters associated with treatments (12.6% of claims-based studies), such as adherence to treatment guidelines [116] and treatments for depression [117].

Claims data were rarely used to study prevention (5.6% of claims-based studies). One of the few studies evaluated the role of provider type (primary care physicians versus endocrinologists) on the receipt of preventive health services by diabetes patients [118], while another examined the effectiveness and costs of influenza vaccinations [119].

Is the Use of Claims Data Increasing?

The use of claims in published healthcare studies differed very little from year-to-year, and there was no continuous up or down trend (see Table **3**; p>.05). This suggests that the small variations from year-to-year were likely the result of unplanned "timing"; that is, the point in time of a submission to a journal, or "revise and resubmit" requests, or the backlog of articles yet-to-be-published in any given year. Indeed, the lack of any trend generally was true for each of the five journals.

We compared the use of claims-based studies in the period 2000-2002 with their use during 2003-2005 in order to determine if the year-to-year data were masking a wider trend. This comparison did not reveal a meaningful increase in the rate of claims-based studies between the periods; 6.5% of the 2000-2002 period were claim-based, as were 8.1% of the studies published in the later period (Phi = .032, p = .154).

Do Researchers Inform Readers of the Weaknesses of the Data?

Only 62 of the 143 studies (43.4%) that used claims data mentioned any problems with the data. Coding errors were the most commonly cited weakness of the data. They were noted by authors of 20 studies, who did not usually specify the type of coding error (for example, whether the coding error was due to potential misinterpretations of the medical record or typographical errors). Rather, the authors simply used the classification, coding error.

The second most frequently mentioned limitation of the data (13 studies) was the lack of detail in the claims, such as information on disease severity [113, 120]. One study, for example, noted that they could not distinguish patients with Type 1 and Type 2 diabetes in claims data [118]. Authors of other studies linked the lack of detail in claims data to problems with analyses. Siran Koroukian and her colleagues [121], for example, noted that additional clinical variables would have allowed the creation of better models.

Selection biases, such as patients choosing an insurer based on their healthcare needs and the type of care they expect to receive [122] or hospital coding practices that cause some diseases to be overrepresented [123], were

Table 3.   Type of Data by Year, Claims *vs*. All Others

			Year						Total
			2000	**2001**	**2002**	**2003**	**2004**	**2005**	
Claims Versus All Others	Not Claims	Count	322	298	291	316	316	270	1813
		% within Year	92.0%	93.7%	95.1%	91.6%	93.2%	90.6%	92.7%
	Claims	Count	28	20	15	29	23	28	143
		% within Year	8.0%	6.3%	4.9%	8.4%	6.8%	9.4%	7.3%
Total		Count	350	318	306	345	339	298	1956
		% within Year	100%	100%	100%	100%	100%	100%	100%

Phi = .055 p = .305.

mentioned by the authors of 6 claims-based studies. Some authors simply stated that different hospitals might have different ways of recording certain hard-to-code conditions. But other authors specifically mentioned upcoding as the cause of diseases being overrepresented. For example, one study noted that "overall accuracy, unreliability, variability in coding practices across facilities, and 'upcoding' of diagnostic information are issues which may also affect model performance" [89].

There was very little difference between the journals with respect to the likelihood of authors of claims-based studies mentioning weaknesses/limitations (p> .05). Studies published in the *American Journal of Medical Quality* were the least likely to state weaknesses/limitations (2 of 7 studies that utilized claims data did mention a limitation). The authors in *The New England Journal of Medicine* did the best job; six of the 10 studies mentioned a limitation (see Table **4**).

There was no statistically significant variation between studies of one healthcare area with claims and those examining another with respect to the likelihood of the authors mentioning weaknesses of the data. This was likely due to the few cases in each category (see Table **5**). Investigators studying treatments and access discussed data weaknesses in less than half their articles (access, 35.7%; treatment, 27.8%). Those who studied the other areas did a better job; more than half of such studies mentioned weaknesses (prevention, 75%; quality, 55.9%; morbidity, 53.8%). Most of the prevention studies noted that claims data lacked detail, which made it impossible to determine whether patients had received a particular procedure [124, 125].

DISCUSSION

The results of this study suggest that the use of claims data in research is small. Overall, claims represented 7.3% of data used in studies published in the five journals. Claims, however, may be a much more common data source among researchers who do not use clinical trials; when the *New England Journal of Medicine* was removed from the analysis, the figure rose to 12.1%. Claims use, however, may have somewhat leveled. The rapid rise in the number of

Table 4.   Journal by Mention of Claims Data Limitations

			Tainted Data Code		Total
			No Mentions	**Mentions**	
Journal Name Code	Medical Care	Count	44	29	73
		% within Journal	60.3%	39.7%	100.0%
	Health Care Financing Review	Count	18	18	36
		% within Journal	50.0%	50.0%	100.0%
	Medical Care Research and Review	Count	10	7	17
		% within Journal	58.8%	41.2%	100.0%
	New England Journal of Medicine	Count	4	6	10
		% within Journal	40.0%	60.0%	100.0%
	American Journal of Medical Quality	Count	5	2	7
		% within Journal	71.4%	28.6%	100.0%
Total		Count	81	62	143
		% within Journals	56.6%	43.4%	100.0%

Phi = .14 p = .588.

Table 5.   Medical Area by Mention of Claims Data Limitations

			Tainted Data Code		Total
			No Mention	Mentions	
Medical Area	Access	Count	45	25	70
		% within Medical Area	64.3%	35.7%	100.0%
	Prevention	Count	2	6	8
		% within Medical Area	25.0%	75.0%	100.0%
	Quality	Count	15	19	34
		% within Medical Area	44.1%	55.9%	100.0%
	Morbidity	Count	6	7	13
		% within Medical Area	46.2%	53.8%	100.0%
	Treatment	Count	13	5	18
		% within Medical Area	72.2%	27.8%	100.0%
Total		Count	81	62	143
		% within Medical Area	56.6%	43.4%	100.0%

Phi = .257 p = .05.

claims-based studies that occurred during the 1990s seems to have stalled.

The largest proportion of claims-based studies (49%) examined issues related to access to healthcare, in particular how changes in healthcare policy affected access [126, 127]. They also examined disparities in treatment [128, 129]. Claims based-research may provide a general idea of the access patients have to healthcare. One study, for example, studied within-hospital racial disparities and compared it to results from studies involving multiple hospitals [128]. The authors concluded that disparities were more likely the result of differences between hospitals than they were the consequence of racial inequality within any single hospital; blacks, in general, went to hospitals that provided poorer quality of care and this explained what appeared to be racial disparity. Such explanations do provide a general portrait of the healthcare situation and are useful. Claims data may be more appropriate for use in access studies and this may explain why the authors of such research were next-to-least likely to mention any of the weaknesses associated with the data; maybe they just did not see the need.

The next largest proportion of claims-based studies (23.8%) examined issues related to quality of healthcare. Claims data may not always be suitable for quality of care studies because they may not be accurate nor detailed enough [13, 29, 30, 111]. These investigators commonly used claims data to evaluate hospital or healthcare provider activities [7, 130]. One study, for example, used claims data to look for inappropriate drug combinations among HIV patients in hospitals [131]. They found that 2% of their units of analysis contained inappropriate drug combinations.[7] Policy decisions, however, should not solely rest on results obtained from such claims-based quality of care studies. In the just noted example, the 2% error rate could easily be the

result of inaccuracies in the data; one study, for example, found that claims differ from the medical record by as much as 27% for prescriptions [132]. A few authors whose studies dealt with the quality of healthcare recognized that the use of claims data was problematic and designed their studies to specifically assess the validity of the data by comparing it to the medical record. All of them found discrepancies between the data sets [e.g., see 62].

Studies of treatments, therapies, and interventions made up 12.6% of the claims-based investigations. Claims data are not uniformly suitable for these studies because procedure coding has been shown to be inaccurate, especially when the procedure is not required for reimbursement [23]. Illustrative, one study in our sample examined whether treatment for alcoholism affected future healthcare utilization [8]. The researchers compared claims data among patients with a recorded diagnosis of alcoholism. Unfortunately, alcoholism has a low level of sensitivity in claims data compared to the medical record, only 20 percent in one study [133], and there is evidence from an older study that the diagnosis is absent in claims data for a substantial portion of individuals who are receiving treatment [134].

There are potential exceptions to the suitability of claims data with respect to studies of treatments, therapies, and interventions. They may be appropriate when used to locate a sample or to establish costs. Illustrative, Angela Sauaia and her colleagues [135] identified acute myocardial infarction (AMI) cases with a fee-for-service database. It is possible that errors or upcoding might result in individuals without an AMI being initially included in such a sample, but further investigation in the just cited study helped insure its integrity; medical chart abstraction was used to confirm the existence of an AMI [135]. The other often acceptable use of claims data is for the establishment of costs, which can be then used to compare the cost effectiveness of procedures (e.g., the cost-effectiveness of new treatments). Arden Morris and her colleagues [17], for example, used claims

---

7 Their unit of analysis was "person-years". They "identified 2110 person-years of claims data representing the claims experience of slightly more than 1000 enrollees in 1999 and 2000" [130: pg III-55].

data to help establish the expenses associated with two competing methods for detecting malignant breast tumors.

Research on morbidity and mortality made up 9.1% of the claims-based studies. Claims data use is problematic in morbidity and mortality studies because diagnosis coding has been shown to be inaccurate [47, 53]. Researchers commonly used the claims data to examine morbidity and mortality rates in different settings. One study, for example, examined re-hospitalization rates and survival rates for patients in health maintenance organizations (HMO) versus fee-for-service plans (FFS) [115; see also 136]. One of the surprising findings from this research illustrates the impact that illegal activities may have on claims data. The researchers noted that HMO stroke patients, despite being healthier, were more likely than their FFS counterparts to be re-hospitalized within 30 days. The result, however, might be due to upcoding. Claims for patients in FFS plans might have been upcoded to obtain additional reimbursement (HMOs have less of an incentive to upcode). The result of the upcoding would be that patients would appear to be sicker in the FFS claims data even if in truth they were not [43]. This would explain the surprising finding of Smith and her colleagues.

Finally, a small proportion (5.6%) of claims-based studies examined issues related to prevention and detection of disease. The studies examined issues such as whether beneficiaries were using preventive health services [121] and whether certain preventive treatments were more cost effective [119]. Claims may be especially unsuitable for these types of evaluation studies because most prevention and detection-based procedures are bundled with other procedures or office visits, or simply not required for reimbursement [21, 22]. Results from these studies must be suspect until the value of the data from which they were derived has proved valid.

In this report we have presented a general picture of the use of claims data in research; in particular we have noted when illegal behavior impacts the data as this is an area that is likely little known among those using the data. Our next step is to review individual studies to assess whether the results are fatally flawed because of weaknesses in the data or whether the data were appropriate for each study. Such an evaluation will assist policymakers to assess the usefulness of individual studies, and will provide healthcare researchers with specifics by which they may gauge the appropriateness of the data for their purposes.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Yan Y, Birman-Deych E, Radford M, Nilasena D, Gage B. Comorbidity indices to predict mortality from medicare data: results from the national registry of atrial fibrillation. Med Care 2005; 43(11): 1073-7.

[2]   Turner B, Newschaffer C, Zhang D, Cosler L, Hauck W. Antiretroviral use and pharmacy-based measurement of adherence in postpartum HIV-infected women. Med Care 2000; 38(9): 911-25.

[3]   Zito J, Foster C, Fox M. Building pharmacoepidemiological capacity to monitor psychotropic drug use among children enrolled in medicaid. Am J Med Qual 2000; 15(4): 126-36.

[4]   Larson M, Miller K, Sharma S, Manderscheid R. Children's mental health services in fee-for-service medicaid. Health Care Financ Rev 2004; 26(1): 5-22.

[5]   Dickey B, Normand S, Drake R, Weiss R, Azeni H, Hanson A. Limiting inpatient substance use treatment: what are the consequences? Med Care Res Rev 2003; 60(4): 332-46.

[6]   National Emphysema Treatment Trial Research Group. Cost effectiveness of lung-volume–reduction surgery for patients with severe emphysema. N Engl J Med 2003; 348(21): 2092-102.

[7]   Ellerbeck E, Engelman K, Williams J, Nazir N, Markello S. Variations in diabetes care and the influence of office systems. Am J Med Qual 2004; 19(1): 12-8.

[8]   Kane R, Wall M, Potthoff S, Stromburg K, Dai Y, Meyer Z. The effect of alcoholism treatment on medical care use. Med Care 2004; 42(4): 395-402.

[9]   Laine C, Lin Y, Hauck W, Turner B. Availability of medical care services in drug treatment clinics associated with lower repeated emergency department use. Med Care 2005; 43(10): 985-95.

[10]   Laliberte L, Fennell M, Papandonatos G. The relationship of membership in research networks to compliance with treatment guidelines for early-stage breast cancer. Med Care 2005; 43(5): 471-9.

[11]   Wheeler J. Can a disease self-management program reduce health care costs? The case of older women with heart disease. Med Care 2003; 41(6): 706-15.

[12]   Nakashima E. Prescription data used to assess consumers: records aid insurers but prompt privacy concerns. Washington Post 2008; Sect.A01.

[13]   Hicks J. The potential of claims data to support the measurement of healthcare quality. Santa Monica, CA: Rand Corporation 2003.

[14]   Krueger D.E. Measurement of prevalence of chronic disease by household interviews and clinical evaluations. Am J Public Health 1957; 47(8): 953-60.

[15]   Baron J, Weiderpass E. An introduction to epidemiological research with medical databases. Ann Epidemiol 2000; 10(4): 200-4.

[16]   Couris C, Schott A, Ecochard R, Morgon E, Colin C. A literature review to assess the use of claims databases in identifying incident cancer cases. health serv outcomes. Res Methodol 2003; 4(1): 49-63.

[17]   Morris A, Flowers C, Morris K, Schmidt W, Pommier R, Vetto J. Comparing the cost-effectiveness of the triple test score to traditional methods for evaluating palpable breast masses. Med Care 2003; 41(8): 962-71.

[18]   Panacek E. Performing chart review studies. Air Med J 2007; 26(5): 206-10.

[19]   Preston A, Chua W, Neu D. The diagnosis-related group prospective payment system and the problem of the government of rationing health care to the elderly. Accoun Org Soc 1997; 22(2): 147-64.

[20]   Centers for Medicare & Medicaid Services. Medicare and medicaid move aggressively to encourage greater patient safety in hospitals and reduce never events [Online]. 2008 Jul 31, [Cited 2008]. Available from: http://www.cms.hhs.gov/apps/media/press/release. asp?Counter=3219&intNumPerPage=10&checkDate=1&checkKey =&srchType=1&numDays=14&srchOpt=0&srchData=&keywordT ype=All&chkNewsType=1%2C+2%2C+3%2C+4%2C+5&intPage =&showAll=1&pYear=&year=0&desc=&cboOrder=date.

[21]   Dresser M, Feingold L, Rosenkranz S, Coltin K. Clinical quality measurement: comparing chart review and automated methodologies. Med Care 1997; 35(6): 539-52.

[22]   Steinwachs D, Stuart M, Scholle S, Starfield B, Fox M, Weiner J. A comparison of ambulatory medicaid claims to medical records: a reliability assessment. Am J Med Qual 1998; 13(2): 63-9.

[23]   Dismuke C. Underreporting of computed tomography and magnetic resonance imaging procedures in inpatient claims data. Med Care 2005; 43(7): 713-17.

[24]   Bright R, Avorn J, Everitt D. Medicaid data as a resource for epidemiologic studies: strengths and weaknesses. J Clin Epidemiol 1989; 42(10): 937-45.

[25]   Benin AL, Vitkauskas G, Thornquist E, *et al.* Validity of using an electronic medical record for assessing quality of care in an outpatient setting. Med Care 2005; 43(7): 691-8.

[26] Quan H, Parsons G, Ghali W. Validity of procedure codes in international classification of diseases, 9th revision, clinical modification administrative data. Med Care 2004; 42(8): 801-9.

[27] Romano P, Mark D. Bias in the coding of hospital discharge data and it implications for quality assessment. Med Care 1994; 32(1): 81-90.

[28] Green J, Wintfield N. How accurate are hospital discharge data for evaluating effectiveness of care? Med Care 1993; 31(8): 719-31.

[29] Romano P. Using administrative data to identify associations between implanted medical devices and chronic diseases. Ann Epidemiol 2000; 10(4): 197-9.

[30] Tyree P, Lind B, Lafferty W. Challenges of using medical insurance claims data for utilization analysis. Am J Med Qual 2006; 21(4): 269-75.

[31] Cooper G, Yuan Z, Stange K, Dennis L, Amini S, Rimm A. Agreement of medicare claims and tumor registry data for assessment of cancer-related treatment. Med Care 2000; 38(4): 411-21.

[32] Pinfold P, Goel V, Sawka C. Quality of hospital discharge and physician data for type of breast cancer surgery. Med Care 2000; 38(1): 99-107.

[33] Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005; 58: 323-37.

[34] United States General Accounting Office. Consultants' billing advice may lead to improperly paid insurance claims [Online Publication No. GAO-01-818]. 2001 Jun, [Cited 2008]. Available from: http://www.gao.gov/new.items/d01818.pdf.

[35] United States Department Of Health And Human Services. office of inspector general. improper fiscal year 2000 medicare fee-for-service payments [Online Publication No. A-17-00-02000]. 2008 Jan 17, [Cited 2008]. Available from: http://oig.hhs.gov/ oas/reports/afma/a0002000.pdf.

[36] United States General Accounting Office. medicare fraud and abuse: DOJ continues to promote compliance with false claims act guidance [Online GAO Publication No. GAO-02-546]. 2002 Apr, [Cited 2008]. Available from: http://www.gao.gov/new.items/ d02546.pdf.

[37] Brown JG. Medicare payments for drg 475: respiratory system diagnosis with ventilator support [Online HHS Publication No. OEI-03-98-00560]. 1998 December, [Cited 2008]. Available from: http://www.oig.hhs.gov/oei/reports/oei-03-98-00560.pdf.

[38] Brown JG. Interoffice memorandum, to: Nancy-Ann Min DeParle, Administrator, Health Care Financing Administration [Online HHS Publication No. OEI-01-98-00420]. 1999 Jan 21, [Cited 2008]. Available from: http://oig.hhs.gov/oei/reports/oei-01-9800420.pdf# search=%92NancyAnn%20Min%20DeParle%20%20January%202 1%2C%%20201999%2C%20%20SUBJECT%3A%20monitoring %20the%20accuracy%20of%20hospital%20coding%20%28OEI01 9800420%29%92.http://tinyurl.com/98bfne.

[39] Hsia D. DRG 121 Validation study update: circulatory disorders with myocardial infarction and cardiovascular complications [Online HHS Publication No. OEI-12-89-00194]. 1992 October, [Cited 2008]. Available from: http://www.oig.hhs.gov/oei/ reports/oei-12-89-00194.pdf.

[40] Rehnquist J. Medicare maintenance payments for capped rental equipment [Online HHS Publication No. OEI-03-00-00410]. 2002 June, [Cited 2008]. Available from: http://www.oig.hhs.gov/oei/ reports/oei-03-00-00410.pdf.

[41] Jesilow P, Pontell HN, Geis G. Prescription for profit: how doctors defraud medicaid. Berkeley: University of California Press 1993.

[42] Sparrow M. License to Steal: how fraud bleeds America's health care system (Updated ed.). Boulder, Colorado: Westview Press 2000.

[43] Jesilow P. The effects of fraud on the evaluation of health care. Health Care Anal 2005; 13(9): 239-45.

[44] Brown JG. A compendium of reports and literature on coding of physician services [Online Publication No. OEI-03-91-00921]. 1994 May, [Cited 2008]. Available from: http://www.oig.hhs.gov/ oei/reports/oei-03-91-00921.pdf.

[45] DOJ Has Improved oversight of false claims act guidance united states general accounting office. Report to congressional committees. March 2001.

[46] Harrington KB. Clamping down on upcoding: government efforts to curb medicare billing fraud and abuse. PhD [unpublished dissertation]. Lincoln (NE): The University of Nebraska 2003.

[47] Psaty B, Boineau R, Kuller L, Luepker R. The potential costs of upcoding for heart failure in the United States. Am J Cardiol 1999; 84(1): 108-9.

[48] Traska M. Unbundling could be costing you a bundle. Bus Health 1990; 8(3): 20-7.

[49] U.S. Department of Health and Human Services and U.S. Department of Justice. Health care fraud and abuse control program annual report for fy 2001 [Online]. 2002 Apr, [Cited 2008]. Available from: http://oig.hhs.gov/reading/hcfac/HCFAC Annual Report FY 2001.htm http://www.oig.hhs.gov/publications/ docs/hcfac/HCFAC%20Annual%20Report%20FY%202001.htm

[50] United States Department of Justice. Health care fraud report: fiscal year 1997 [Online]. [Cited 2008]. Available from: http://www. usdoj.gov/dag/pubdoc/health97.htm#execsum.

[51] Brown JG. Medicare Payments for DRG 014: specific cerebrovascular disorders except transient ischemic attack [Online Publication No. OEI-03-99-00240]. 1999 July, [cited 2008]. Available from: http://209.85.173.132/search?q=cache:sBdChVmT74cJ:www.oig.h hs.gov/oei/reports/oei-03-99-00240.pdf+Medicare+Payments+for+ DRG+014:+Specific+Cerebrovascular+Disorders+Except+Transie nt+Ischemic+Attack&hl=en&ct=clnk&cd=1&gl=us.

[52] Lu M. Separating the true effect from gaming in incentive-based contracts in health care. JEMS 1999; 8(3): 383-431.

[53] Steele L, Glazier R, Lin E, Evans M. Using administrative data to measure ambulatory mental health service provision in primary care. Med Care 2004; 42(10): 960-5.

[54] Klabunde C, Harlan L, Warren J. Data sources for measuring comorbidity: a comparison of hospital records and medicare claims for cancer patients. Med Care 2006; 44(10): 921-8.

[55] Lee D, Donovan L, Austin P, *et al*. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. Med Care 2005; 43(2): 182-8.

[56] Kwon A, Bungay K, Pei Y, *et al*. Antidepressant use: concordance between self-report and claims records. Med Care 2003; 41(3): 368-74.

[57] Du X, Freeman J, Warren J, Nattinger A, Zhang D, Goodwin J. Accuracy and completeness of medicare claims data for surgical treatment of breast cancer. Med Care 2000; 38(9): 719-27.

[58] McCarthy E, Iezzoni L, Davis R, *et al*. Does clinical evidence support icd-9-cm diagnosis coding of complications? Med Care 2000; 38(8): 868-76.

[59] Weingart S, Iezzoni L, Davis R, *et al*. Use of administrative data to find substandard care: validation of the complications screening program. Med Care 2000; 38(8): 796-806.

[60] Birman-Deych E, Waterman A, Yan Y, Nilasena D, Radford M, Gage B. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Med Care 2005; 43(5): 480-5.

[61] Penberthy L, McClish D, Manning C, Retchin S, Smith T. The added value of claims for cancer surveillance: results of varying case definitions. Med Care 2005; 43(7): 705-12.

[62] Aronsky D, Haug P, Lagor C, Dean N. Accuracy of administrative data for identifying patients with pneumonia. Am J Med Qual 2005; 20(6): 319-28.

[63] Ahmed F, Janes G, Baron R, Latts L. Preferred provider organization claims showed high predictive value but missed substantial proportion of adults with high-risk conditions. J Clin Epidemiol 2005; 58(6): 624-8.

[64] Fisher ES, Whaley FS, Kmushat WM, *et al*. The accuracy of medicare's hospital claims data: progress has been made, but problems remain. Am J Public Health 1992; 82(2): 243-8.

[65] Blumstein A, Cohen J, Nagin D, Eds. Deterrence and incapacitation: estimating the effects of criminal sanctions on crime rates. Washington: National Academy of Sciences 1978.

[66] SCI-BYTES: The ten most cited journals of 2006 [Online]. 2007 Sept 10, [Cited 2007]. Available from: http://www.in-cites.com/ research/2007/september_10_2007-1.html.

[67] The New England Journal of Medicine: Research & review articles on diseases and clinical practice [Online]. [Cited 2007]. Available from: http://content.nejm.org/.

[68] American College of Medical Quality Homepage [Online]. [Cited 2007]. Available from: http://www.acmq.org/index.cfm.

[69] SAGE Publications – American Journal of Medical Quality [Online]. [Cited 2007]. Available from: http://www.sagepub.com/ journalsProdDesc.nav?prodId=Journal201749.

[70]   Medical Care Home [Online]. [Cited 2007]. Available from: http://www.lww-medicalcare.com/pt/re/medcare/home.htm.

[71]   SAGE Publications – Medical Care Research and Review [Online]. [Cited 2007]. Available from: http://www.sagepub.com/journals ProdDesc.nav?prodId=Journal200970.

[72]   Health Care Financing Review Overview [Online]. [Cited 2007]. Available from: http://www.cms.hhs.gov/HealthCareFinancing Review/.

[73]   Siderowf A. Evidence from clinical trials: can we do better? Neurotherapeutics 2004; 1(3): 363-71.

[74]   Everitt B, Pickles A. Statistical aspects of the design and analysis of clinical trials. Singapore: World Scientific Publishing 2004.

[75]   Friedman L. Fundamentals of clinical trials. New York: Springer Science+Business Media 1999.

[76]   Patel V, Kushniruk A, Yang S, Yale J. Impact of a computer-based patient record system on data collection, Knowledge Organization, and Reasoning. JAMIA 2000; 7(6): 569-85.

[77]   Willis J. Medical records: use and abuse. Br J General Pract 2004; 54(506): 725.

[78]   Lloyd S, Rissing P. Physician and coding errors in patient records. JAMA 1985; 254(10): 1330-6.

[79]   Jones L, Clarke W, Carney C. Receipt of diabetes services by insured adults with and without claims for mental disorders. Med Care 2004; 42(12): 1167-75.

[80]   Woods C. Impact of different definitions on estimates of accuracy of the diagnosis data in a clinical database. JCE 2001; 54(8): 782-8.

[81]   Luck J, Peabody J, Dresselhaus T, Lee M, Glassman P. How well does chart abstraction measure quality? a prospective comparison of standardized patients with the medical record. Am J Med 2000; 108(8): 642-9.

[82]   Bogardus ST Jr, Geist DE, Bradley EH. Physicians' interactions with third-party payers: is deception necessary? Arch Intern Med 2004; 164(7): 1841-4.

[83]   Novack DH, Detering BJ, Arnold R, Forrow L, Ladinsky M, Pezzullo JC. Physicians' attitudes toward using deception to resolve difficult ethical problems. JAMA 1989; 261(20): 2980-5.

[84]   Rost K, Smith R, Matthews DB, Guise B. The deliberate misdiagnosis of major depression in primary Care. Arch Fam Med 1994; 3(4): 333-7.

[85]   Werner RM, Alexander GC, Fagerlin A, Ubel PA. The "hassle factor": what motivates physicians to manipulate reimbursement rules? Arch Intern Med 2002; 162(10): 1134-9.

[86]   Wynia MK, Cummins DS, Vangeest JB, Wilson IB. Physician manipulation of reimbursement rules for patients: between a rock and a hard place. JAMA 2000; 283(14): 1858-65.

[87]   Freeman V, Rathore S, Weinfurt K, Schulman K, Sulmasy D. Lying for patients: physician deception of third-party payers. Arch Intern Med 1999; 159(19): 2263-70.

[88]   Betz ME, Gebo KA, Barber E, *et al.* Admissions in a multistate cohort of HIV-positive adults in 2001. Med Care 2005; 43(9): 3-14.

[89]   Sloan K, Sales A, Liu CF, *et al.* Construction and characteristics of the rxrisk-v: a VA-adapted pharmacy-based case-mix instrument. Med Care 2003; 41(6): 761-74.

[90]   Rosen A, Loveland S, Anderson J, *et al.* Evaluating diagnosis-based case-mix measures: how well do they apply to the VA population? Med Care 2001; 39(7): 692-704.

[91]   Austin PC, Tu JV, Alter DA, Naylor CD. The impact of under coding of cardiac severity and comorbid diseases on the accuracy of hospital report cards. Med Care 2005; 43(8): 801-9.

[92]   Berg K, Mor V, Morris J, Murphy K, Moore T, Harris, Y. Identification and evaluation of existing nursing homes quality indicators. Health Care Financ Rev 2002; 23(4): 20-36.

[93]   Romano P, Schembri M, Rainwater J. Can administrative data be used to ascertain clinically significant postoperative complications? Am J Med Qual 2002; 17(4): 145-54.

[94]   Romano P, Yasmeen S, Schembri M, Keyzer J, Gilbert W. Coding of perineal lacerations and other complications of obstetric care in hospital discharge data. Obstet Gynecol 2005; 106(4): 717-25.

[95]   Lofthus C, Cappelen I, Osnes E, *et al.* Local and national electronic databases in Norway demonstrate a varying degree of validity. JCE 2005; 58(3): 280-5.

[96]   Frankfort-Nachmias C, Nachmias D. Research methods for the social sciences, 6th ed. New York: Worth Publishers 2000.

[97]   Bowling A. Research methods in health: investigating health and health services. Buckingham: Open University Press 2002.

[98]   Dooley D. Social research methods. Upper Saddle River, New Jersey: Prentice Hall 2001.

[99]   Orne MT. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. Am Psychol 1962; 17: 776-83.

[100]   Britten N. Qualitative interviews. qualitative research. In: Pope C, Mays N, Eds. Health Care, Malden: Blackwell Publishing Inc 2006; pp. 12-20.

[101]   Dendukuri N, McCusker J, Bellavance F, *et al.* Comparing the validity of different sources of information on emergency department visits: a latent class analysis. Med Care 2005; 43(3): 266-75.

[102]   Hylek E, Go A, Chang Y, *et al.* Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. NEJM 2003; 349(11): 1019-26.

[103]   Jick T. Mixing quantitative and qualitative methods: triangulation in action. Adm Sci Q 1979; 24(4): 602-11.

[104]   Shavers V, Brown M, Klabunde C, *et al.* Race/Ethnicity and the intensity of medical monitoring under 'watchful waiting' for prostate cancer. Med Care 2004; 42(3): 239-50.

[105]   Bradford WD, Kaste LM, Nietert PJ. Continuity of medical care, health insurance, and nonmedical advice in the first 3 years of life. Med Care 2004; 42(1): 91-8.

[106]   Haber S, Khatutsky G, Mitchell J. Covering uninsured adults through medicaid: lessons from the Oregon health plan. Health Care Financ Rev 2000; 22(2): 1-17.

[107]   Cotter J, Smith W, Rossiter L, Pugh C, Bramble J. Combining state administrative databases and provider records to assess the quality of care for children enrolled in medicaid. Am J Med Qual 1999; 14(2): 98-104.

[108]   Fowles J, Fowler E, Craft C, *et al.* Comparing claims data and self-reported data with the medical record for pap smear rates. Eval Health Prof 1997; 20(324): 324-42.

[109]   Kressin N, Chang B, Hendricks A, Kazis L. Agreement between administrative data and patients' self-reports of race and ethnicity. Am J Public Health 2003; 93(10): 1734-9.

[110]   Boehmer U, Kressin N, Berlowitz D, Christiansen C, Kazis L, Jones J. Self-Reported *vs.* administrative race/ethnicity data and study results. Am J Public Health 2002; 92(9): 1471-3.

[111]   Kieszak S, Flanders W, Kosinski A, Shipp C, Karp H. A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. JCE 1999; 52(2): 137-42.

[112]   Steele L, Glazier R, Lin E, Austin P, Mustard C. Measuring the effect of a large reduction in welfare payments on mental health service use in welfare-dependent neighborhoods. Med Care 2005; 43(9): 885-91.

[113]   Eppig F, Poisal J. Medicare beneficiaries' use of prescription drug discount cards, CY 2002. Health Care Financ Rev 2003; 25(2): 91-4.

[114]   Wu B, Pope G. Left ventricular ejection fraction test rates for medicare beneficiaries with heart failure. Am J Med Qual 2002; 17(2): 61-6.

[115]   Smith M, Frytak J, Liou J, Finch M. Rehospitalization and survival for stroke patients in managed care and traditional medicare plans. Med Care 2005; 43(9): 902-10.

[116]   Sokol M, McGuigan K, Verbrugge M, Epstein R. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005; 43(6): 521-30.

[117]   Donohue J, Berndt E, Rosenthal M, Epstein A, Frank R. Effects of pharmaceutical promotion on adherence to treatment guidelines for depression. Med Care 2004; 42(12): 1176-85.

[118]   Lafata E, Martin M, Morlock R, Divine G, Xi H. Provider type and the receipt of general and diabetes-related preventive health services among patients with diabetes. Med Care 2001a; 39(4): 361-72.

[119]   Davis J, Lee E, Taira D, Chung R. Influenza vaccination, hospitalizations, and costs among members of a medicare managed care plan. Med Care 2001; 39(12): 1273-80.

[120]   Burge F, Lawson B, Johnston G. Family physician continuity of care and emergency department use in end-of-life cancer care. Med Care 2003; 41(8): 992-1001.

[121]   Koroukian S, Litaker D, Dor A, Cooper G. Use of preventive services by medicare fee-for-service beneficiaries: does spillover from managed care matter? Med Care 2005; 43(5): 445-52.

[122]   Spetz J, Smith M, Ennis S. Physician incentives and the timing of cesarean sections: evidence from California. Med Care 2001; 39(6): 536-50.

[123]   Groeneveld P, Laufer S, Gerber A. Technology diffusion, hospital variation, and racial disparities among elderly medicare beneficiaries: 1989-2000. Med Care 2005; 43(4): 320-9.

[124]   Brennan P, Kagay C, Geppert J, Moos R. Predictors and outcomes of outpatient mental health care: a 4-year prospective study of elderly medicare patients with substance use disorders. Med Care 2001; 39(1): 39-49.

[125]   Lafata E, Cole C, Ben-Menachem T, Morlock R. Sociodemographic differences in the receipt of colorectal cancer surveillance care following treatment with curative intent. Med Care 2001b; 39(4): 361-72.

[126]   Kenney G, Sommers A, Dubay L. Moving to mandatory medicaid managed care in Ohio: impacts on pregnant women and infants. Med Care 2005; 43(7): 683-90.

[127]   LoSasso A, Freund D. A longitudinal evaluation of the effect of medical managed care on supplemental security income and aid to families with dependent children enrollees in two California counties. Med Care 2000; 38(9): 937-47.

[128]   Barnato A, Lucas FL, Staiger D, Wennberg D, Chandra A. Hospital-level racial disparities in acute myocardial infarction treatment and outcomes. Med Care 2005; 43(4): 308-19.

[129]   Epstein A, Weissman J, Schneider E, Gatsonis C, Leape L, Piana R. Race and gender disparities in rates of cardiac revascularization: do they reflect appropriate use of procedures or problems in quality of care? Med Care 2003; 41(2): 1240-55.

[130]   Downey L, Lafferty W, Tao G, Irwin K. Evaluating the quality of sexual health care provided to adolescents in medicaid managed care: a comparison of two data sources. Am J Med Qual 2004; 19(1): 2-11.

[131]   Hellinger F, Encinosa W. Inappropriate drug combinations among privately insured patients with HIV disease. Med Care 2005; 43(9): 53-62.

[132]   Tamblyn R, Lavoie G, Petrella L, Monette J. The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Quebec. JCE 1995; 48(8): 999-1009.

[133]   Fowles JB, Fowler EJ, Craft C. Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. J Ambul Care Manage 1998; 21(1): 24-34.

[134]   Schwartz AH, Perlman BB, Paris M, Schmidt K, Thornton JC. Psychiatric diagnoses as reported to medicaid and as recorded in patient charts. Am J Public Health 1980; 70(4): 406-8.

[135]   Sauaia A, Ralston D, Schluter W, Marciniak T, Havranek E, Dunn T. Influencing care in acute myocardial infarction: a randomized trial comparing 2 types of intervention. Am J Med Qual 2000; 15(5): 197-206.

[136]   Yuan Z, Cooper G, Einstadter D, Cebul R, Rimm A. The association between hospital type and mortality and length of stay: a study of 16.9 million hospitalized medicare beneficiaries. Med Care 2000; 38(2): 231-45.