CURRENT OPINION

Pharmacoeconomics 1999 Jul; 16 (1): 1-8
1170-7690/99/0007-0001/$04.00/0

Using Healthcare Claims Data for Outcomes Research and Pharmacoeconomic Analyses

Howard G. Birnbaum,[1] *Pierre Y. Cremieux*,[1] *Paul E. Greenberg*,[1] *Jacques LeLorier*,[2] *JoAnn Ostrander*[1] and *Laura Venditti*[1]

1 Analysis Group/Economics, Cambridge, Massachusetts, USA
2 Centre de Recherche, Centre Hospitalier de l'Universite de Montreal, Montreal, Quebec, Canada

Abstract

Healthcare claims data are a practical complement to data from randomised controlled trials (RCTs) for evaluating health outcomes in non-experimental settings and for generalising results to a broader population. Claims data are a relatively inexpensive way to obtain useful information about patient demographics, as well as healthcare resources used for specific medical conditions and procedures from large numbers of patients over extended periods of time. With claims data, it is possible to identify patients who meet specific medical or sociodemographic criteria, estimate their costs, define episodes of medical care, and measure outcomes more globally than is possible with RCT data.

Statistical methods exist to address some of the inherent issues with claims data due to their limited clinical detail. We also identify extensions of claims data to productivity issues, the use of centralised claims data such as in Canada, and the application of new statistical methods to outcomes research literature such as sample selection correction methods.

Consumers of health outcomes research, such as third-party payers (TPPs), pharmaceutical companies and researchers, increasingly require accurate, clinically credible, statistically valid and economically meaningful information about the outcomes and effectiveness of alternative medical treatments and services. For example, Canada and Australia have developed economic evaluation criteria for use as part of a mandatory approval process for new prescription drug products.[1]

Pharmacoeconomics combines the use of economics, epidemiology, decision analysis and biostatistics into a comprehensive evaluation of treatments.[2] The estimation of healthcare costs for a specific patient population is integral to any contemporary pharmacoeconomic investigation, whether it be from a cost-effectiveness, cost-benefit or cost-utility approach.[3] Cost-effectiveness methods involve the comparison of intervention costs in monetary terms with intervention effectiveness measured in clinical units. Using claims data may allow estimation of both the cost numerator and outcomes-measure denominator. Which costs and outcomes are included in the numerator and denominator are dependent on the perspective of the analysis (i.e., from the patient, payer or societal point of view).

Traditionally, pharmaceutical-related health outcomes research has relied upon medical records and/or patient-derived data, such as case report forms from randomised controlled trials (RCTs) for new medications. Limitations of RCTs often

include their lack of detailed economic data, questionable external validity in non-experimental settings, extensive time and resources required, and in many cases, limited sample size.[4] RCT data are typically collected on a narrowly targeted population that meets specific clinically relevant criteria and follows carefully defined treatment protocols. While such data are rich in clinical detail, they do not reflect real world experience regarding costs and utilisation of healthcare services that is especially meaningful to health policy discussions about such matters as coverage of new pharmaceuticals. While detailed health economic outcomes data are increasingly collected in RCTs, the data are only collected under experimental rather than usual care conditions.[5] Healthcare claims data (i.e. that covering medical and prescription services and products) rectify many of the problems of RCT data, although they have their own shortcomings.

Healthcare benefit claims are a practical complement to RCT data for evaluating health outcomes in usual care settings and generalising results to a broader population. While a RCT may provide data to measure efficacy, RCT data are extremely unlikely to define effectiveness in the real world. Healthcare claims data would seem to be a practical way to evaluate health outcomes in non-experimental settings and to generalise results to a broader population. Government and private sector healthcare claims are important data sources for health outcomes research because they provide economic information on actual medical practice. For example, Mitchell[6] reviewed the use of Medicare hospital and physician claims data for outcomes research by 10 Patient Outcomes Research Teams (PORTs) supported by the US Agency for Healthcare Policy Research (AHCPR).[7] All of the AHCPR-funded PORTs included components that used insurance claims databases to examine outcomes. Similarly, LeLorier et al.[8] discussed outcomes research using claims data from the Quebec Universal health-insurance-plan database, *Regie de l'Assurance-Maladie de Quebec* (RAMQ).

Outcomes research literature has focused on treatment evaluation, quality of care and other considerations in a wide range of therapeutic areas (e.g. low birth-weight babies,[9] stroke,[10] pneumonia,[11] cancer[12] and cataracts[13]). As prescription claims data have become more widely available, their use for outcomes research has grown. Prescription drug claims data can be used to address such pharmacoeconomic research questions as:

- What is the impact of receiving a particular type of antidepressant on depressed patients' medical costs?[14]
- What is the impact of under-utilisation of medication treatment in a particular group of patients?[15]
- Is the use of particular pharmaceuticals a risk factor for glaucoma?[16]

Claims data are collected on a routine basis for administrative purposes (i.e. paying providers for services delivered to beneficiaries). Hence, they provide relatively inexpensive information, from large numbers of patients over long time periods, on patient demographics, as well as healthcare resources used for specific medical conditions and procedures. This information on the provision of healthcare services to individuals can be linked over time, across settings and by providers.

Statistical methods exist to address some of the inherent issues with claims data due to their limited clinical detail.[6,17-19] We identify extensions of claims data to productivity issues, the use of centralised claims data such as in Canada, and the application of new statistical methods to outcomes research literature such as sample selection correction methods.

1. Claims Data Availability and Applicability

Outcomes researchers have made extensive use of government claims data files for the Medicare and Medicaid programmes in the US, which are primarily for elderly and low income beneficiaries, respectively, as well as employer claims data. The federal Medicare programme maintains the most accessible and analytically developed databases, and encourages the use of its analytic files devel-

oped specifically for health economics and outcomes researchers.[20] However, while hospital, physician and outpatient services are covered under Medicare, the programme generally does not cover outpatient pharmaceutical use. Consequently, prescription drug claims, together with non-reimbursed expenses, are not available from Medicare. An important exception involves Medicare coverage of recombinant human erythropoietin (rHuEPO) administered to patients with end-stage renal failure.[21]

State operated Medicaid programmes provide insurance coverage to low income, and other 'categorically or medically needy' populations, many of whom are children. Medicaid coverage includes prescription drug benefits, as well as hospital, physician and related services. Consequently, Medicaid medical and prescription claims data have been used extensively in pharmacoeconomic research.[22,23] However, while these data are available for outcomes research, the results may not be easily generalised and their use has been limited because Medicaid patients are much more likely than the population as a whole to be poor, chronically ill and/or institutionalised.[24] A further limitation is that neither government nor private sector prescription claims data indicate the type of drugs given to inpatients.[25]

In contrast to the limited availability of prescription drug data from Medicare and the lack of generalisability of the Medicaid data, private benefit plans often include prescription drug coverage for employees, dependents and retirees. As a result, private plan medical and prescription claims data are widely used in outcomes research, and are increasingly used in pharmacoeconomic analyses. While private healthcare claims data are not readily available, special arrangements can be made with particular payers to use their data for specific outcomes research purposes.[26] Various vendors also provide private claims databases for research purposes.[27]

Some TPPs cover prescription drugs as part of a major medical benefit. However, with the advent of managed care, TPPs are increasingly using prescription benefit managers (PBMs) to administer this benefit. Prescription data available through PBMs generally contain considerably more detailed information than are available through a TPP's major medical programme. PBM programmes track a beneficiary's identification number, as well as the specific drugs, dosage and number of days for which the prescription is dispensed. (These data are typically available using a National Drug Code (NDC) which identifies the specific labeller, product and package size codes. Also, the therapeutic class codes, dosage form and strength and a generic indicator are often available). Moreover, these prescription claims data can be linked through the beneficiary's identifier with medical claims files containing administrative details on patients' encounters with healthcare providers. In the absence of such managed pharmaceutical programmes, prescription claims data, if available at all, are usually limited to counts of prescriptions dispensed, and often omit specific details, such as medication name, dosage and price.

As healthcare claims data analysis unfolds, its potential to be linked to other emerging sources of related data, such as disease registries, absenteeism from work, and reduced productivity while at work grows. For example, ongoing pharmacoeconomics research by the authors uses the health care and disability claims data of patients with specific diseases and matched controls to analyse employer direct and indirect costs. Similar research has been conducted by others.[28] The Surveillance, Epidemiology and End Results (SEER)/Medicare database developed by the National Cancer Institute and Health Care Financing Administration (HCFA) combines specific population-based registry data with Medicare claims files, permitting the study of associations between various cancer patient and healthcare setting factors and treatment patterns, costs and medical outcomes.[29,30] These types of opportunities have the potential to make claims data even more useful.

For example, one potentially useful type of data that is not available from traditional healthcare claims databases is information concerning indirect costs to employers. For some diseases such as depression, indirect costs can dwarf direct costs

typically measured with traditional data approaches.[31] For analyses focused on these issues, work-related indirect cost data may be available from the employer which can be linked with healthcare claims.[32] These data include sporadic absenteeism from work, short or long term disability, workers compensation and reduced productivity while at work.

Centralised healthcare claims data in Canada have been used to conduct comprehensive evaluations of overall payer costs. For example, 2 Canadian provinces, Quebec and Saskatchewan, have created comprehensive databases to administer their single-payer healthcare systems. The Saskatchewan databases cover all the medical and pharmaceutical benefits received by practically all the approximately 1 million inhabitants of the province. Consequently, these databases have an important advantage of providing extremely useful, large, population-based denominators. Another asset of the Saskatchewan databases is the possibility of linking the data with hospital chart and discharge summaries. Thus, it is not surprising that several widely disseminated papers have emerged from these databases.[33-35] The use of the Quebec databases for outcomes research is more recent but has also produced important information on the ocular adverse effects of inhaled steroids,[16,36] the impact of ciprofloxacin in the rates of hospitalisation for chronic bronchitis,[37] the management of new symptoms of dyspepsia in Quebec elderly[38] and the use of lipid-lowering agents.[23]

2. Issues Involved in the Appropriate Use of Claims Data

Healthcare claims data allow for rapid and inexpensive tracking of the overall medical experiences of large populations and subpopulations of patients using multiple providers.[39] By constructing analytical files that gather patient-specific information from multiple claims, such databases provide a single comprehensive source of data on patient care, provider services, resource utilisation and organisational settings. It is then possible to identify patients who meet specific study criteria (e.g. patients under the age of 18 years with asthma), estimate costs for these patients (in the form of third-party expenditures), define episodes of medical care and measure outcomes of treatment.[5] In addition, it is possible to use claims data to overcome potential RCT sample selection bias problems.

2.1 Identify Patients

To ensure the analysis of healthcare claims data of comparable persons and conditions, it is critical to develop a clinically meaningful set of selection criteria with both inclusion and exclusion criteria (e.g. documentation of a particular diagnosis or a prescription for a particular type of drug during a certain period). Frequently, diagnostic and procedure codes are used to select initial patient populations; subsequently, some persons are eliminated from the sample who do not meet certain additional clinical and/or enrolment exclusionary criteria, such as no antidepressant use for 6 months prior to filling an index prescription, or no substance abuse diagnoses.[14]

2.2 Estimate Costs

Unlike RCTs, which typically focus on the use of 1 drug or intervention, healthcare claims data allow a more global, longitudinal understanding of the implication of a particular drug treatment or medical intervention. For example, from claims or RCT data, it is possible to identify the impact of a switch among antidepressants on the probability of hospitalisation. However, only claims data can identify actual third-party expenditures for the hospitalisation, as well as other related healthcare services, and additionally provide the opportunity to track these data over extended durations. From claims data, it is also possible to identify expenditures not only from the payer's perspective but also from that of the employer by using disability and absenteeism data often linked to healthcare utilisation claims. It is then possible to estimate the relative costs of various interventions from a global perspective and include potentially important indirect costs. Direct healthcare costs are commonly estimated based on payments or amount reim-

bursed, or less frequently and less desirably using billed or allowed charges, or a microcosting technique. Methods to extract actual costs include the use of department or institution-specific cost-to-charge ratios.[40]

2.3 Define Episodes

To ensure analysis of clinically meaningful patterns of treatment, it is helpful to use the dates as well as diagnostic and procedural information on the claims to define an episode of care. The episode refers to the period commencing with the initial claim for a particular treatment under study, or index prescription, and ending when no treatment has occurred for a meaningful period of time.[26] For example, in the case of antidepressants, Frank et al.[41] used an 'eight-week period without treatment to separate treatment episodes.' Because the data are 'censored' by the start and end dates of the claims files, it is important to leave 'windows' of time at the start- and end-points to eliminate potentially censored episodes from the analytical file. Often, expert clinical guidance is integral in defining episode 'rules' because of varying lengths of treatment for different illnesses. Because claims databases can track patient care over long periods of time, they are particularly useful in longitudinal studies of chronic diseases.[42]

2.4 Measure Outcomes

An important challenge in using healthcare claims data is the measurement of outcomes. While some clinical outcomes cannot be measured with claims data (e.g. visual acuity or psychosocial functioning), these data can be used to measure outcomes such as the percentage of preterm births, hospital mortality rates and complication rates.[6] For example, researchers can estimate the probability that a patient with particular characteristics was hospitalised or had an adverse health outcome such as an emergency room visit associated with home infusion therapy.[17]

Claims databases are well suited for measuring outcomes and other phenomena that essentially reflect analytical manipulation of dates and types of service. For example, Mennemeyer and Winkelman[39] used claims data to detect clinical laboratory errors. While the data measured for such analyses are straightforward, the outcomes measured by these data are potentially powerful and important.

2.5 Overcome Potential RCT Sample Selection Bias Problems

As information technology improves, some patients have become sophisticated users of clinical trials as alternative ways of obtaining treatment. As a result, even in placebo-controlled trials, results might not accurately reflect potential outcomes when applied to a population at large that did not volunteer for the trial. Using claims files allows analysis of a truly population-based sample of insured patients. In contrast, placebo-controlled trials require that patients agree to participate and sign informed consent forms. These RCT procedures can generate concerns regarding over-representation of patients without health insurance, with more severe illnesses and with lower socioeconomic demographics, than the population at large.

3. Issues Associated with Claims Data Use

In contrast to RCT data, the main shortcoming of healthcare claims data for pharmacoeconomic research is the limited amount of information they contain regarding clinical severity, patient health status and process of care. Healthcare claims data include the patient's diagnoses, occurrence of a procedure, and details of services provided during treatment. Because claims data lack clinical details, they may be vague about a patient's particular outcome or specific treatment.[18,43] Despite these limitations, researchers have used information based on claims data to account, in part, for severity of illness using measures including:

- comorbid diagnoses (e.g. number of and presence of specific diagnoses on a claim)
- prior use of healthcare services (e.g. previous hospitalisations)
- frequency and types of medications.

Although validity problems can affect all types of research, including RCTs, they appear in distinctive ways in healthcare claims data research. One threat to validity is incomplete assembly of longitudinal data (e.g. missing bills and/or multiple plan coverage). For example, in the US, claims data on health maintenance organisation (HMO) enrollees are often unavailable (e.g. if the plan is capitated and encounter data are not available). In contrast, in the Canadian provincial model, this concern is virtually eliminated. As a result, comprehensive databases such as the RAMQ in Quebec have become quite popular for pharmacoeconomic research.[7,34,35]

Because healthcare claims data are primarily structured to provide for third-party (i.e. insurance) payments to providers, they contain less than optimal clinical information for outcomes research. Some epidemiologists consider the clinical imprecision of claims data to be a fatal flaw; however, other investigators have demonstrated the validity of identifying people with specific diseases, such as diabetes.[44]

Not only may clinical details be limited, but healthcare claims data may also include erroneous diagnostic and procedural coding. For example, there may be incentives to 'upcode' diagnostic codes when reimbursement is related to diagnoses and/or procedures, as in Medicare's Diagnostic Related Groups (DRGs) system, i.e. payments are made according to severity or level of care.[18] Other limitations also may exist. For example, the claims may not document all services provided because claim forms may limit the number of diagnoses that can be documented. Alternatively, some managed health plans that use staff physicians do not receive such information while other more traditional indemnity plans may not record details of prescription claims paid under a major medical benefit plan.

Indication bias (or confounding by indication) is another challenge faced by investigators using administrative databases for outcomes research. It occurs when patients are targeted to receive an intervention based on the prognosis of the intervention under study. For instance, a new drug that is perceived to be safer might be prescribed to the most vulnerable patients with a resulting higher rate of adverse events. Thus, it can be difficult to determine whether higher rates of adverse events are due to the inherent characteristics of the drug (high toxicity) or to the patients to whom it was prescribed (high vulnerability). This bias is usually controlled by means of statistical methods such as multiple regression analyses. The problem, however, might remain when using claims data, due to inadequate capture of important confounders.

Fortunately, there are approaches to address many aspects of the inherent limitations of healthcare claims data for pharmacoeconomic analysis.[45] Since claims data provide usage information in usual care settings, as opposed to the controlled settings of clinical trials, they can be especially useful in moving beyond efficacy to measure the effectiveness of specific treatments. Ultimately, the decision to use claims data for outcomes research depends on the appropriateness of claims data and the availability, cost and suitability of alternative data sources to answer these questions.[46]

As claims data become increasingly available, statistical tools designed for their analysis are emerging in the outcomes research literature. For example, sample selection correction methods used in the economics evaluation literature are increasingly relied upon in outcomes and pharmacoeconomic research to address issues of nonrandom attrition and missing data.[19] New approaches to the treatment of missing data that are particularly well suited to the analysis of longitudinal data typical of administrative claims are also becoming more accessible. For example, software to apply multiple imputation methods to incomplete data is now available.[47,48] In addition, econometric tools developed to analyse situations in which no control group is available for comparison lend themselves naturally to many real world situations in which pharmacoeconomic investigations can rely on health claims data.[49]

4. Conclusion

When seeking data for health outcomes research, the critical first step is to define the data elements necessary to attain a rigorous answer to the proposed research question. If the research involves economic issues or clinical encounter questions related to particular diagnoses or procedures, the use of health claims databases is a likely source of data.

Claims data have their limitations, but when used appropriately they can provide a powerful tool to fill what otherwise would be a void regarding detailed economic information and generalisable conclusions. For example, it may be possible to estimate the cost of a particular disease for the economy as a whole rather than only for the healthcare sector. Employer costs resulting from a disease or avoided by a new treatment can be precisely estimated from claims data.

As improvements are made to claims databases such as the standardisation of claims information reporting, the addition of pertinent clinical variables, and inclusion of information relative to patient severity of illness and quality of life, opportunities will expand for conducting health outcomes research using claims data. Recent Canadian research using provincial claims data offer a good example of what can be done. A natural extension of the existing body of healthcare claims data research is to conduct integrated studies that combine claims data for participants in RCTs with clinical data derived from the trial.[26] Such studies would address growing concerns over potential bias in RCTs resulting from increased patient awareness of the potential benefits of using pharmaceuticals in these trials.

While previous literature identifies some of the issues discussed here, recent research broadens the applications of claims data and addresses particular methodological issues. For example, recent research in the US illustrates that healthcare data can be linked analytically to patient-specific productivity indicators such as disability or at-work performance data. Pharmacoeconomic analyses using linked healthcare and productivity data can provide useful information on both the direct (i.e. medical treatment) and indirect (e.g. work loss days due to disability) costs to employers for their beneficiaries with specific diseases. Recent research in Canada demonstrates how it is possible to use centralised claims data to conduct a comprehensive evaluation of overall payer costs associated with a specific therapy that was previously difficult to evaluate precisely. Finally, new statistical methods have been implemented to address sample selection bias issues in claims data analyses. The increased flow of information available to patients enrolled in RCTs introduces new concerns about sample selection biases that can be avoided through claims data analysis.

Acknowledgements

The views expressed here are those of the authors alone, and do not necessarily reflect the opinions of the institutions with which they are affiliated.

References

1. Canadian Coordinating Office for Health Technology Assessment (CCOHTA). Guidelines for economic evaluation of pharmaceuticals: Canada. 2nd ed. Ottawa (ON): CCOHTA, 1997
2. Schulman KA, Linas BP. Pharmacoeconomics: state of the art in 1997. Annu Rev Public Health 1997; 18: 529-48
3. Weinstein MC, Siegel JE, Gold MR, et al. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. JAMA 1996; 15: 1253-8
4. Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges and strategies. Clin Ther 1997; 19 (2): 346-66
5. Lewis NJW, Patwell JT, Briesacher BA. The role of insurance claims databases in drug therapy outcomes research. Pharmacoeconomics 1993; 4 (5): 323-9
6. Mitchell JB. Using Medicare claims for outcomes research. Med Care 1994; 32 (7 Suppl.): JS38-51
7. Agency for Health Care Policy and Research. Outcomes of pharmaceutical therapy: program update [online]. Available from: URL: http://www.ahcpr.gov/clinic/pharmout.htm. [accessed 1998 Oct 16]
8. LeLorier J, Poirier S, Page V, et al. The effect of a $2 co-payment on prescription refill rates of Quebec elderly and its relationship to socio-economic status. Can Pharm J 1998; 130 (10): 30-4
9. Goldenberg RL. When infection causes pre-term labor. OBG Manage 1997; Jun: 55-68
10. Stukenborg GJ. Comparison of carotid endarterectomy outcomes from randomized controlled trials and Medicare administrative databases. Arch Neurol 1997; 54: 826-32
11. Metlay JP, Schulz R, Li Y-W, et al. Influence of age on symptoms at presentation in patients with community-acquired pneumonia. Arch Intern Med 1997; 157: 1453-9
12. Weeks J, Pfister DG. Outcomes research studies. Oncology 1996; 10 (11 Suppl.): 29-34

13. Bass EB, Wills S, Scott IU, et al. Preference values for visual states in patient planning to undergo cataract surgery. Med Decis Making 1997; 17: 324-30
14. Russell JM, Berndt ER, Miceli R, et al. Course and cost of treatment with SSRI's [abstract NR585]. Poster Session, New Research Program and Abstracts. American Psychiatric Association, 1998 Annual Meeting; 1998 Jun 3; Toronto
15. Harrow BS, Strom BL, Gans JA. Impact of pharmaceutical underutilization: a study of insurance drug claims data. J Am Pharm Assoc 1997; NS37 (5): 511-6
16. Garbe E, LeLorier J, Boivin JF, et al. Inhaled and nasal glucocorticoids and the risk of ocular hypertension and open-angle glaucoma. JAMA 1997; 277: 722-7
17. Birnbaum HG, Tang M. The home infusion therapy/relative benefit index (HIT/RBI): summary of an analysis using insurance claims data. Med Care 1998; 36 (5): 757-65
18. Iezzoni LI. Risk adjustment for measuring health care outcomes. Ann Arbor (MI): Health Administration Press, 1994
19. Crown B, Obenchain R, Englehart L, et al. The application of sample selection models outcomes research: the case of evaluating the effects of antidepressant therapy on resource utilization. Stat Med 1998; 17: 1943-58
20. HCFA data users reference guide [online]. Available from: URL: www.hcfa.gov/stats [accessed 1998 Oct 7]
21. Griffiths RI, Powe NR, Greer J, et al. A review of the first year of medicare coverage of erythropoietin. Health Care Financ Rev 1994; 15 (3); 83-102
22. Monane M, Bohn RL, Avorn J, et al. Compliance with antihypertensive therapy among elderly Medicaid enrollees: the roles of age, gender and race. Am J Public Health 1996; 86 (12): 1805-8
23. Avorn J, Monette J, LeLorier J, et al. Persistence of use of lipid-lowering medications: a cross-national study. JAMA 1998; 279 (18): 1458-62
24. Bright RA, Avorn J, Everitt DE. Medicaid data as a resource for epidemiologic studies: strengths and limitations. J Clin Epidemiol 1989; 42 (10); 937-45
25. Davis H. Use of computerized health claims data to monitor compliance with antibiotic prophylaxis in sickle cell disease. Pharmacoepidemiol Drug Saf 1997; 7: 107-12
26. Thompson D, Buesching D, Gregor KJ, et al. Patterns of antidepressant use and their relation to costs of care. Am J Managed Care 1996; 2 (9): 1239-46
27. Topol EJ, Ellis SG, Cosgrove DM, et al. Analysis of coronary angioplasty practice in the United States with an insurance-claims data base. Circulation 1993; 87 (5): 1489-97
28. Greenberg P, Bailit H, Berndt ER, et al. Illness and productivity: objective workplace evidence [working paper]. Boston (MA): MIT Program on the Pharmaceutical Industry, 1997
29. US Department of Health and Human Services, AHCPR. SEER-Medicare linked databases: a compendium of selected published health data sources. New York: Walcoff & Associates, 1996: 87-8
30. Riley GF, Potosley AL, Lubitz JD, et al. Medicare payments from stage at diagnosis. Med Care 1995; 33 (8): 828-41
31. Greenberg PE, Stiglin LE, Finkelstein SN, et al. The economic burden of depression in 1990. J Clin Psychiatry 1993; 54 (11): 405-26
32. Greenberg P, Birnbaum H, Barton M, et al. Tackling costs one disease at a time. Business & Health 1999; 17 (5): 31-2, 37
33. Spitzer WO, Suissa S, Ernst P, et al. The use of β-agonists and the risk of death and near death from asthma. N Engl J Med 1992; 326: 501-5
34. Blais L, Ernst P, Boivin JF, et al. Inhaled corticosteroids and the prevention of readmission to hospital for asthma. Am J Respir Crit Care Med 1998; 158: 126-32
35. Ray WA, Griffin MR, West R, et al. Incidence of hip fracture in Saskatchewan, Canada, 1976-1985. Am J Epidemiol 1990; 131 (3): 502-9
36. Garbe E, Lelorier J, Suissa S. Risk of occular hypertension of open angle glaucoma in elderly patients on oral glucocorticoid. Lancet 1997; 350: 979-82
37. LeLorier J, Derderian F. The effect of listing ciprofloxacin in provincial formularies on hospitalizations for bronchitis and pyelonephritis. Can J Clin Pharmacol 1998; 5 (3): 133-7
38. LeLorier J, Page V, Castilloux AM, et al. The management of new symptoms of dyspepsia in Quebec elderly. Can J Gastroenterol 1997; 11 (8): 669-72
39. Mennemeyer ST, Winkelman JW. Searching for inaccuracy in clinical laboratory testing using Medicare data. JAMA 1993; 269 (8): 1030-3
40. Lave, JR, Pashos CL, Anderson GR, et al. Costing medical care: using Medicare administrative data. Med Care 1994; 32 (7): JS77-89
41. Frank RG, Busch SH, Berndt ER. Measuring prices and quantities of treatment for depression. AEA Papers Proc 1998; 88 (2): 106-11
42. Grady ML. Medical effectiveness research data methods. Rockville (MD): Agency for Health Care Policy and Research (AHCPR), 1992. AHCPR publication no.: 92-0056
43. Hannan EL, Racz MJ, Jollis JG, et al. Using Medicare claims data to assess provider quality for CABG surgery: does it work well enough? Health Serv Res 1997; 31 (6): 659-78
44. Young TK, Roos NP, Hammerstrand KM. Estimated burden of diabetes mellitus in Manitoba according to health insurance claims: a pilot study. Can Med Assoc J 1991; 144 (3): 318-24
45. The Robert Wood Johnson Foundation. Working with large insurance databases: avoiding and overcoming the pitfalls. Washington, DC: The Robert Wood Johnson Foundation's Changes in Health care Financing and Organization Program, Alpha Center, 1995
46. US General Accounting Office. Regulatory reform: information on costs, cost-effectiveness, and mandated deadlines for regulations. Briefing report to the ranking minority member. Washington, DC: Committee on Governmental Affairs, US Senate, 1995
47. Statistical Solutions. Solas for missing data analysis: the solution for missing values in your data. User reference manual. Cork, Ireland: Statistical Solutions, 1997
48. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Stat 1978; 6 (1): 34-58
49. Mennemeyer ST. Can econometrics rescue epidemiology. Ann Epidemiol 1997; 7 (4): 249-50

Correspondence and reprints: Dr *Howard G. Birnbaum*, Analysis Group/Economics, One Brattle Square, Cambridge MA 02138, USA.
E-mail: hbirnbaum@ag-inc.com