Concordance between childhood injury diagnoses from two sources: an injury surveillance system and a physician billing claims database

A Kostylova, B Swaine, D Feldman

....

See end of article for authors' affiliations

....

Correspondence to:
Professor B Swaine, Centre de recherche interdisciplinaire en réadaptation du Montréal métropolitain (CRIR), Institut de réadaptation de Montréal, 6300, avenue Darlington, Montréal, Québec, Canada, H3S 2J4; bonnie.swaine@ umontreal.ca

Accepted 9 January 2005
....

Objectives: (1) To determine the concordance between injury diagnoses (head injury (HI), probable HI, or orthopedic injury) for children visiting an emergency department for an injury using two data sources: an injury surveillance system (Canadian Hospitals Injury Research and Prevention Program, CHIRPP) and a physician billing claims database (Régie de l'assurance maladie de Québec, RAMQ), and (2) to determine the sensitivity and specificity of diagnostic and procedure codes in billing claims for identifying HI and orthopedic injury among children.

Design: In this cross sectional cohort, data for 3049 children who sought care for an injury (2000–01) were obtained from both sources and linked using the child's personal health insurance number.

Methods: The physician recorded diagnostic codes from CHIRPP were used to categorize the children into three groups (HI, probable HI, and orthopedic), while an algorithm, using ICD-9-CM diagnostic and procedures codes from the RAMQ, was used to classify children into the same three groups.

Results: Concordance between the data sources was ''substantial'' (weighted Kappa 0.66; 95% CI 0.63 to 0.69). The sensitivity of diagnostic and procedure codes in the RAMQ database for identifying HI and for orthopedic injury were 0.61 (95% CI 0.57 to 0.64) and 0.97 (95% CI 0.96 to 0.98), respectively. The specificity for identifying HI and for orthopedic injury were 0.97 (95% CI 0.96 to 0.98) and 0.58 (95% CI 0.56 to 0.63), respectively.

Conclusion: Combining diagnostic and procedures codes in a physician billing claims database (the RAMQ database) may be a valid method of estimating injury occurrence among children.

Information about pediatric injury in Canada can be gleaned from a number of sources. These include population surveys (for example, the National Longitudinal Survey of Children and Youth and the National Population Health Survey), medical record reviews, hospital and trauma registries (including administrative databases), and injury surveillance systems.[1] With regards to the latter, the Canadian Hospitals Injury Research and Prevention Program (CHIRPP) is a primary source of information on patterns of childhood injury in Canada.

CHIRPP is a computerized emergency room based injury surveillance system that operates in 10 pediatric and five general hospitals across the country. It gathers important data (such as mechanism of injury, nature of injury, and body part) relating to children's visits to hospital emergency departments for injury. Pickett and collaborators have provided an overview of the system's strengths and weaknesses.[2] Its strengths are: (1) information on the circumstances in which injuries occur that cannot be obtained elsewhere, (2) its high rate of parental compliance, and (3) its data are invaluable for the development of appropriate preventive interventions. Weaknesses of CHIRPP include its need for active cooperation of doctors and emergency staff as well as technical support (adding to the expense of running such a system), and its representativeness, as the rate of children's use of emergency departments after sustaining an injury could vary among communities.[3–5] Finally, it is somewhat limited in that it does not provide data relating to injuries for which care was sought outside the emergency department (such as pediatrician's or physician's office).

Physician billing claims represent another data source for injury research. Taylor *et al*, used the billing records (fee-for-service billing data) of a tertiary care pediatric emergency department in Halifax, Nova Scotia, Canada to predict the number of future emergency department visits for trauma compared with those for respiratory infection.[6] They concluded that billing data, that includes diagnostic codes based on the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM),[7] provide a good measure of future trauma occurrence risk among children. In Quebec, the fee-for-service database of the Quebec Health Insurance Board, or the Régie de l'assurance maladie du Québec (RAMQ), was shown to provide a valid source of information for outcome assessment of injuries in the elderly.[8] The sensitivity of the claims data for the measurement of *any injury* was found to be 81.3%, but the sensitivities for *specific injuries* varied widely from 14.3% to 97.2%. The accuracy of this type of data was also found to be high in cases of cardiovascular prognosis,[9] hip fracture,[10–12] and stroke.[13] Although these data are primarily used for reimbursement and accounting purposes, physician billings claims have been suggested as an appropriate data source for quality of care assessment,[14–16] pharmacoepidemiologic research,[17] technology assessment,[18] evaluation of medical care appropriateness,[19] and cost utilization studies.[20] The advantages of using claims data are: (1) large samples of geographically dispersed patients, (2) longitudinal records, (3) convenience and low cost (the data are already collected and available), and (4) defined sampling frames.[14 21–25] Limitations of these types of administrative data include inaccurate coding that is thought to be due to patient complexity, vague context of ICD-9-CM code definitions, institution-specific variations in

....

Abbreviations: CHIRPP, Canadian Hospitals Injury Research and Prevention Program; HI, head injury; RAMQ, Régie de l'assurance maladie de Québec.

coding practices, and financial incentives to record certain diagnoses.[9] [26] [27] To our knowledge, the use of physician claims data has never been formally validated for pediatric injury, and in particular for head injury (HI). Moreover, physician claims data have never been validated against an injury surveillance system.

The purpose of this study was to compare data gathered using the CHIRPP surveillance system with that from the physician claims files from the RAMQ (Quebec Provincial Health Insurance Board) for the same group of children who visited an emergency department for the same index injury. Specifically, we sought to determine (1) the agreement between childhood injury diagnoses (HI and orthopedic injury), and (2) the sensitivity and specificity of diagnostic and procedure codes in billing claims for identifying HI and orthopedic injury among children.

METHODS

Data from the CHIRPP database were compared with those from the physician claims database of the RAMQ. CHIRPP data were collected at the emergency departments of the two pediatric hospitals in Montreal from December 2000 to October 2001 (Montreal Children's Hospital and Hôpital Saint-Justine). The CHIRPP database contains information about the circumstances of the injury that is completed by the patient or an accompanying adult on a one page self administered CHIRPP questionnaire. Also on the back of the same form, the physicians record the following clinical information: diagnostic codes, injured body part, and treatment provided. The child's personal health insurance number (a 10 digit number that provides a unique identification number for each beneficiary of the Quebec health insurance plan), date of birth, sex, postal code, and date of visit are also recorded on the CHIRPP sheet.

The Quebec Health Insurance Board or RAMQ—the government agency responsible for administering the provincial health insurance plan—uses a computerized billing service to record service use and to reimburse service providers. This database contains the physician's identification number, the patient's provincial health insurance number, the date and location of visit, type of consultation, as well as the ICD-9-CM diagnostic codes for the visit, procedure codes for services provided, and the cost of these services.

Procedure

The records of 3145 children aged 1–18 years who sought care for an injury during the study period were identified and extracted from the CHIRPP database. Children under one year of age were excluded because they have not received their permanent health insurance number and we would be unable to link them with the RAMQ database.

The physician recorded diagnostic codes in the CHIRPP database separated the children into three groups: (HI, probable HI, and orthopedic). This classification was chosen for the purposes of a larger study examining the risk of second HI among children. The HI group was defined as children with a diagnosis of HI (for example, skull fracture, minor HI, concussion, intracranial injury, and multiple injuries with associated HI). When a child sustained multiple injuries and one of the injuries included an HI, the child was classified into the HI group. The probable HI group included children with injuries to the eyes, face, and teeth or isolated facial lacerations only when accompanied with one of the following mechanisms of injury: struck forcefully against a hard surface, a fall from a height, or both. The orthopedic group consisted of children diagnosed with an orthopedic injury of the upper or lower extremity (for example, fracture, sprain, and dislocation).

Table 1 ICD-9CM diagnostic and procedure codes associated with injury

	Definition of code
Diagnostic codes	
800–804	Fracture (skull and face)
850–8549	Concussion
872–874	Laceration (skull and face)
830	Subluxation
910	Trauma to face
920	Contusion of eyes
959	Unspecified trauma
9083	Trauma to cerebral vessels
810–819	Fracture (upper extremity)
820–829	Fracture (lower extremity)
880–884	Open wound of upper extremity
890–894	Open wound of lower extremity
Procedure codes	
1320	Simple laceration (face and neck)
2113	Incision, drainage of skull
2505–2527	Treatment for trauma to skull or face
7500–7507	Treatment for skull fracture
7595–7598	Treatment for laceration of skull
8010–8030	X ray (skull or face)
8258–8259	Tomography head
8570	MRI (head or neck)

A file containing this CHIRPP based information (diagnostic group, age, sex, postal code) was sent by registered mail to the statistical department of the RAMQ (Service des statistiques of the RAMQ). It was returned to us containing the complete registry of services paid to fee-for-service physicians who provided care for 3049 children during 12 months after the index visit for an injury. Data were unattainable for 96 children due to missing RAMQ numbers. The confidentiality of subjects was maintained using a scrambled personal insurance number. This however prevented any additional linking of the two data sources.

To enable a comparison between the two data sources with respect to injury diagnosis, the RAMQ data were grouped into the same three diagnostic categories. This was done using an algorithm (formed of ICD-9-CM diagnostic and procedure codes) developed and validated by a team of researchers (including an emergency department pediatrician) with a sample of 500 children. The HI group was defined as all children who received health services for an HI (ICD-9-CM codes 800, 801, 803, 804, 850–8549, 9083 or procedure codes specific to HI 7500–7507, 7595–7598) or either one of these codes. The probable HI group included all children who had a combination of the following diagnostic (ICD-9-CM 802, 830, 873, 910, 920, 959) and procedure codes (1320, 2113, 2505–2527, 8010–8030 8258–59, 8570). For example, a child having a visit billed with a diagnostic code for concussion (ICD-9-CM 8540) was assumed to have had an HI. A child with a diagnostic code of imprecise trauma (8290) and a procedure code indicating magnetic resonance imaging of the head (8570) was assumed to have had a probable HI. The orthopedic group consisted of all children using health services for an injury to the upper or lower extremities (ICD-9-CM 810–817, 820–829, and 831–839). A list of the pertinent diagnostic and procedure codes associated with injury are presented in table 1.

Analysis

Two analyses were performed. Firstly, we created a 3×3 table (HI, probable HI, and orthopedic) and calculated the percent overall agreement and a weighted kappa statistic between childhood injury diagnoses obtained from the two data sources. Secondly, using CHIRPP as a gold standard, we calculated sensitivity and specificity for HI and orthopedic injury by creating a 2×2 table where the HI group was

Table 2 Characteristics of children seeking care at two Montreal pediatric trauma centers, 2000–01

	HI, n = 724	Probable HI, n = 423	Orthopedic, n = 1902	Total, n = 3049
Sex				
Female	251 (35%)	139 (33%)	815 (43%)	1205 (34%)
Male	473 (65%)	284 (67%)	1087 (57%)	1844 (66%)
Age (years)				
0–4	251 (34%)	189 (45%)	343 (18%)	783 (26%)
5–9	225 (31%)	157 (37%)	460 (24%)	842 (28%)
10–14	164 (23%)	62 (15%)	771 (41%)	997 (33%)
15–18	84 (12%)	15 (3%)	328 (17%)	427 (14%)
Diagnosis				
Superficial injury		25 (6%)	1	26 (0.8%)
Open wound		398 (94%)	32 (2%)	430 (14%)
Fracture	19 (3%)		1047 (55%)	1066 (35%)
Dislocation			154 (8%)	154 (5%)
Strain or sprain			668 (35%)	668 (22%)
Minor head injury	633 (87%)			633 (21%)
Concussion	71 (10%)			71 (2.2%)
Intracranial injury	1			1
Where injury occurred				
Home	311 (42%)	276 (65%)	594 (31%)	1181 (39%)
School	152 (21%)	38 (9%)	454 (24%)	644 (21%)
Park	66 (9%)	25 (6%)	282 (15%)	373 (12%)
Recreation center	86 (12%)	21 (5%)	275 (14%)	382 (13%)
Public place	61 (8%)	23 (5%)	204 (11%)	288 (9%)
Day care	13 (2%)	19 (4%)	28 (1%)	60 (2%)
Other	35 (5%)	21 (5%)	65 (3%)	121 (4%)
Treatment				
Advice only	217 (30%)	7 (1.6%)	29 (1%)	253 (8%)
Treated, follow up	425 (59%)	414 (98%)	1765 (93%)	2604 (85%)
Short stay ED	26 (4%)	0 (0%)	56 (3%)	82 (3%)
Admitted	55 (8%)	2 (0.4%)	51 (3%)	108 (4%)
Missing data	1		1	2

Data collected as part of the Canadian Hospitals Injury Research and Prevention Program (CHIRPP).

combined with the probable HI group and compared to the orthopedic group. Sensitivity for HI was defined as the probability of having an HI (HI group and probable HI group) indicated in the RAMQ database given that HI was recorded as the CHIRPP diagnosis. Specificity was defined as the probability of not having an HI recorded in the RAMQ database when HI was truly absent as indicated by the CHIRPP diagnoses.

RESULTS
The descriptive characteristics of the sample are presented in table 2. There were more males (60.5%) than females with an injury, and across all groups, injuries mostly occurred at home (39%), followed by those at school inside and outside (21%) and in recreation centers (13%). Thirty percent of

children with HI were sent home from the emergency department with advice only, whereas 98% with probable HI and 93% of those with orthopedic injury were treated before being sent home.

Concordance between the two data sources was found to be ''substantial' (weighted Kappa 0.66; 95% CI 0.63 to 0.69) according to the interpretation scale of Landis and Koch, and percent agreement was 81%.[28]

There were 446 exact matches for HI, 173 for probable HI, and 1849 for orthopedic injury. Discordant cases (n = 581 children) were individually examined to determine the underlying reasons of disagreement. Reasons for disagreement varied and some were more easily accounted for than others (table 3). For example, among the discordant cases, 52% were billed with the diagnostic code ''multiple

Table 3 Examination of discordant childhood injury cases based on physician billing records* and pediatric injury surveillance† data (n = 581 children)

Reasons for discordance	Frequency	%
Billed with diagnostic code ''multiple unspecified wound or trauma'' (95% were classified by CHIRPP as HI or probable HI)	304	52
Billed with diagnostic codes indicating HI or probable HI (41% were classified by CHIRPP as orthopedic)	99	17
Billed with a diagnosis unrelated to injury (for example, sinusitis, chicken pox) (96% were classified by CHIRPP as HI or probable HI)	107	18
Billed as having an UE or LE fracture (84% were classified by CHIRPP as HI or probable HI)	44	8
Billed with a diagnosis possibly related to HI (for example, virus infection, nausea, headache) (100% were classified by CHIRPP as HI or probable HI)	11	2
Lacked a procedure or a diagnostic code precluding classification into one of three groups	16	3

*Physician billing claims database (Régie de l'assurance maladie de Québec, RAMQ).
†Canadian Hospitals Injury Research and Prevention Program (CHIRPP).

unspecified wound or trauma'' and 95% of these (n = 289) were classified by CHIRPP as HI or probable HI. The mean age of children among the discordant pairs (6.6 years) was significantly different (p<0.001) from that of the rest of sample (8.5 years).

The sensitivity of diagnostic and procedure codes in the RAMQ database for identifying HI and for orthopedic injury were 0.61 (95% CI 0.57 to 0.64) and 0.97 (95% CI 0.96 to 0.98), respectively. The specificity for identifying HI and for orthopedic injury were 0.97 (95% CI 0.96 to 0.98) and 0.58 (95% CI 0.56 to 0.63), respectively.

DISCUSSION

We compared childhood injury diagnoses using CHIRPP data with that from a physician claims database for the same group of children who visited an emergency department for the same index injury. Our results indicate that the concordance between the two data sources is ''substantial'' and that the sensitivity of claims data for identifying orthopedic injury was higher than that for identifying HI. There are several possible reasons for the less than optimal level of concordance. Firstly, we observed a high frequency of non-specific diagnostic codes in physician billings. It is interesting to note that of the 304 cases that were billed for ''multiple unspecified wound or trauma'', 95% (n = 289) of these were classified by CHIRPP as HI or probable HI. Another 18% of cases (n = 107) were billed with a diagnosis unrelated to injury (for example, sinusitis, chicken pox) yet 96% of these had HI or probable HI CHIRPP based diagnoses. It would thus appear that the physician claims database underestimates the frequency of HI (and probable HI) by approximately 12% (that is, (107+289)/3049). Our results also suggest that the concordance may be associated with the child's age because the children in the discordant pairs were significantly younger than the concordant pairs. This is probably related to a less precise diagnosis among younger children.

Other possible reasons for our results include coding errors in both databases. Clearly, there are errors when completing RAMQ reimbursement forms because 107 children among the discordant pairs, who completed a CHIRPP form *for an injury*, were billed with a RAMQ diagnostic code *other than trauma*. This phenomenon is however unlikely to be related to fee structure (that is, financial incentives). Reimbursement fees for an HI are not necessarily higher than those for an orthopedic injury but appear to vary in both cases according to the time of day of the visit, day of the week of the visit, and so on. Errors could also have occurred in the transcribing of physician notes onto RAMQ reimbursement forms. Certain recording errors of injuries using the CHIRPP system may also exist because 99 children among the discordant pairs were billed with diagnostic codes indicating HI or probable HI but classified by CHIRPP as an orthopedic injury. Perhaps these children suffered multiple injuries and for some reason the physician coded the visit as one for an HI. Physicians who write the initial diagnoses do not necessarily complete the CHIRPP form—sometimes this information is recorded by a clerk using the child's medical file as reference. Finally, a small percentage of missing diagnostic codes in the RAMQ data may also have contributed to the less than perfect concordance.

Indeed, each database was developed for a specific purpose and has its own strengths and weaknesses. Although CHIRPP provides important detailed information about injuries treated in an emergency department, one must keep in mind that the injury surveillance system is not population based. This is in contrast to the physician claims database used in this study, which covers the full continuum from ambulatory to hospital based care, and provides information on almost all

contacts with physicians in the healthcare system. Physician claims data are however limited to healthcare systems where fee-for-services payment is the predominant means of reimbursement. This is the case in Quebec and other parts of Canada, but it may differ in other countries where physicians are paid by capitation or salary. Clearly, there is no single best source for monitoring injury in a population. It would be ideal to link these datasets (injury data with administrative data) to complement the unique strengths of each type of data and provide a more complete picture of childhood injury.

To our knowledge this is the first study in Quebec that measures the accuracy of physician claims data for pediatric injury including HI. We were interested in knowing if physician claims data could be used for another purpose other than physician reimbursement. We believe this study demonstrates a new application and the potential capabilities of using diagnostic and procedures codes from physician billing claims data to study injury (including HI) among children. In particular, one could determine with relative confidence the number of children who receive medical care for an HI, where they receive this care, and the costs associated with emergency department visits for HI.

A number of limitations should be considered in the interpretation of the results. This study involved data for children in Montreal, Quebec and the results may not be generalizable to older populations or to those in other regions. Ideally, sensitivity and specificity estimates require that subjects be classified into diagnostic groups using an error-free gold standard. As discussed above, misclassification within the gold standard could have occurred.

In Quebec, where fee-for-service billing is the predominant method of remuneration, the combination of diagnostic and procedures codes in a physician claims database may be a valid method of estimating injury occurrence among children. Its use may however lead to an underestimation of the frequency of visits for HI, particularly in younger children.

. . . .

Authors' affiliations
A Kostylova, B R Swaine, D Feldman, Centre de recherche interdisciplinaire en réadaptation du Montréal métropolitain, Institut de réadaptation de Montréal; Montreal, Quebec, Canada

---

Key points

- This is the first study to compare surveillance data with physician claims data in the context of pediatric head injury (HI).
- The concordance between the two data sources was ''substantial'' (weighted Kappa 0.66; 95% CI 0.63 to 0.69).
- The sensitivity of claims data for identifying orthopedic injury (0.97, 95% CI 0.96 to 0.98) was higher than that for identifying HI (0.61, 95% CI 0.57 to 0.64).
- Possible reasons for the less than optimal level of concordance include a high frequency of non-specific diagnostic codes in physician billings (52% were multiple unspecified wound or trauma) and coding errors in both data sources.
- The combination of diagnostic and procedures codes in a physician claims database may be a valid method of estimating injury occurrence among children, but may underestimate the frequency of visits for HI, particularly in younger children.

REFERENCES

1 **Pless IB**, Millar W. Unintentional injuries in childhood: results from Canadian health surveys. Ottawa: Health Canada, 2000.
2 **Pickett W**, Brison RJ, Mackenzie SG, *et al.* Youth injury data in the Canadian Hospitals Injury Reporting and Prevention Program: do they represent the Canadian experience? *Inj Prev* 2000;**6**:9–15.
3 **Macarthur C**, Pless IB. Sensitivity and representativeness of a childhood injury surveillance system. *Inj Prev* 1999;**5**:214–16.
4 **Stone DH**, Morrison A, Smith GS. Emergency department injury surveillance systems: the best use of limited resources? *Inj Prev* 1999;**5**:166–67.
5 **Mackenzie SG**, Pless IB. CHIRPP: Canada's principal injury surveillance program. Canadian Hospitals Injury Reporting and Prevention Program. *Inj Prev* 1999;**5**:208–13.
6 **Taylor BW**, MacIntyre J, Forgeron PR. Trauma recurrence in the pediatric emergency population. *J Trauma* 1999;**46**:479–82.
7 International classification of diseases, 9th revision (clinical modification). 5th edn. Washington: US Department of Health and Human Services; 1996, Cat no 86-72897.
8 **Tamblyn R**, Reid T, Mayo N, *et al.* Using medical services claims to assess injuries in the elderly: sensitivity of diagnostic and procedure codes for injury ascertainment. *Am J Epidemiol* 2000;**53**:183–94.
9 **Jollis JG**, Ancukiewicz M, DeLong ER, *et al.* Discordance of databases designed for claims payment versus clinical information systems. Implications for outcomes research. *Ann Intern Med* 1993;**119**:844–50.
10 **Levy AR**, Mayo NE, Grimard G. Rates of transcervical and pertrochanteric hip fractures in the province of Quebec, Canada, 1981–1992. *Am J Epidemiol* 1995;**142**:428–36.
11 **Ray WA**, Griffin MR, Adams ML. Identification of fractures from computerized Medicare files. *J Clin Epidem* 1992;**45**:703–14.
12 **Demlo LK**, Campbell PM, Brown SS. Reliability of information abstracted from patients' medical records. *Med Care* 1978;**16**:995–1005.
13 **Mayo NE**, Danys I, Carlton J, *et al.* Accuracy of hospital discharge coding for stroke. *Can J Cardiol* 1993;**9**(suppl.D):121D–124D.
14 **Lohr KN**. Use of insurance claims data in measuring quality of care. *Int J Technol Assess Health Care* 1990;**6**:263–71.
15 **Ross NP**, Wennberg JE, Malenka DJ, *et al.* Mortality and reoperation after open and transurethral resection of the prostate for benign prostatic hyperplasia. *N Engl J Med* 1989;**320**:1120–4.
16 **Weiner JP**, Powe NR, Steinwachs DM, *et al.* Applying insurance claims data to assess quality of care: a complication of potential indicators. *QRB* 1990;**16**:424–38.
17 **Ray WA**, Griffin MR. Use of Medicaid data for pharmacoepidemiology. *Am J Epidemiol* 1989;**129**:837–49.
18 **Steinberg EP**, Whittle J, Anderson GF. Impact of claims data research on clinical practice. *Intl J of Tech Assess Health Care* 1990;**6**:282–7.
19 **Kosecoff J**, Chassin MR, Fink A, *et al.* Obtaining clinical data on the appropriateness of medical care in community practice. *JAMA* 1987;**258**:2538–42.
20 **Hillman BJ**, Joseph CA, Mabry MR, *et al.* Frequency and cost of diagnostic imaging in office practice—a comparison of self-referring and radiologist-referring physicians. *N Engl J Med* 1990;**323**:1604–8.
21 **Roos LL**, Mustard CA, Nicol JP, *et al.* Registries and administrative data: organization and accuracy. *Med Care* 1993;**31**:201–12.
22 **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**:318–24.
23 **Quam L**, Ellis LB, Venus P, *et al.* Using claims data for epidemiologic research. The concordance of claims-based criteria with the medical record and patient survey for identifying a hypertensive population. *Med Care* 1993;**31**:498–507.
24 **Colditz GA**, Martin P, Stampfer MJ, *et al.* Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am J Epidemiol* 1986;**123**:894–900.
25 **Paganini-Hill A**, Ross RK. Reliability of recall of drug usage and other health-related information. *Am J Epidemiol* 1982;**116**:114–22.
26 **Fisher ES**, Whaley FS, Krushat WM, *et al.* The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health* 1992;**82**:243–8.
27 **Iezzoni LI**. Using administrative diagnostic data to assess the quality of hospital care: Pitfalls and potential of ICD-9-CM. *Int J Tech Assess Health Care* 1990;**6**:272–81.
28 **Landis JR**, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.