Australian hospital data: not just for funding

Jude L Michel and Terri J Jackson

Abstract
Collections of routine, or ‘administrative’, hospital data have many applications in health care and are now recognised as valuable sources of information. In recent decades, administrative data have been seen primarily as funding and billing tools to assist with the reimbursement of hospitals for services provided; this purpose remains the primary focus of the clinical coder workforce. More recently, hospital data have been recognised as valuable resources for a range of health system improvement processes beyond funding. The focus of this paper is to review and demonstrate the diverse uses of administrative data in health services research and quality improvement. By gaining an understanding of how the data are used, we can appreciate the importance of good quality data from the perspective of its multiple uses. This paper describes a sample of the studies conducted in Australia using administrative data in health care improvement.

Keywords (MeSH):
ICD-10-AM; Data Collection; Medical Records; Statistical Data Interpretation; Australia

Patient safety and quality
Administrative data are increasingly being used to detect and monitor hospital-acquired conditions. Early studies using this data found it difficult to detect patient complications, due to the absence of information on the timing of onset of a condition (Hargreaves 2001), and had to rely on codes which reflect complications of care within the code descriptor itself (‘T80-88’ or end of chapter codes in ICD-10-AM). The introduction of the condition onset flag nationwide, in July 2008, enhances routine data, making it more effective in its support of patient safety.

Prior to its introduction, studies had demonstrated the usefulness of the condition onset flag using Victoria’s ‘prefixed’ data. Victoria has assigned a prefix to codes to reflect the timing of onset (as well as whether the diagnosis was a ‘primary’ or ‘associated’ diagnosis) for over 20 years. One of the first reported uses of this data element was that by Carroll, McLean and Walsh (2003), using the ‘C’ prefix to monitor adverse events in a Melbourne teaching hospital. Though this study utilised the condition onset flag, it only included codes specifically associated with a complication of surgical or medical care, rather than the full range of patient adverse events.

A later study by Jackson and colleagues, using the entire Victorian Admitted Episodes Database for 2000/01, identified that 8% of all admissions, and 12% of multi-day admissions, had at least one C-prefixed (‘not present on admission’) diagnosis. They also found that 41% of these prefixed diagnoses were codes not associated with a ‘T’ or ‘Y’ coded complication of care (Jackson et al. 2006). Thus, limiting patient safety efforts to complications identifiable using the ‘T’, ‘Y’ or end of chapter codes would identify only 59% of hospital-acquired conditions.

Other studies have used admitted patient data to analyse adverse events in smaller populations such as lung cancer patients (Cantsilieris, Jackson & Street 2006) and elective surgery patients (Moje, Jackson & McNair 2006). Both these studies utilised the Victorian ‘C’ or complication prefix to identify the whole range of adverse events experienced by these patient groups. Cantsilieris and colleagues (2006) found that lung cancer patients had predictably high rates of complications (23% of episodes), with the highest
rates recorded for surgical and admitted radiotherapy episodes. Moje and colleagues (2006) found unexpectedly high rates of complications amongst elective surgery patients (15.5%). Using surgical code blocks in the Australian Classification of Health Interventions, they demonstrated very high complication rates associated with CABG (67% of all episodes), colectomy (52%), and hip and knee arthroplasty (42% and 36% respectively).

Following the Bundaberg Hospital scandal, Queensland Health increased its focus on monitoring health outcomes by developing a technique using variable life-adjusted display (VLAD) control charts (Duckett, Coory & Sketcher-Baker 2007). This entails using the Queensland Hospital Admitted Patient Data Collection (QHAPDC) to monitor 31 indicators on a monthly basis. VLAD indicators are flagged when the number of adverse outcomes exceeds the risk-adjusted state average, with thresholds established for notifying more senior levels of the health system the more a hospital’s performance deviates from the expected rate.

Investigation is focused on patterns in the data, rather than individual cases in order to determine whether the detected variation is due to data quality, substandard patient care or differences in casemix; whatever the reason, it is important to identify the underlying cause (Duckett et al. 2007). While it is not possible to risk-adjust the routine hospital data to account for all factors known to affect outcomes, further investigation following a flag allows the hospital to submit recoded data, if appropriate, when the likely cause for a decline in performance is unrecorded changes in their casemix. One benefit of using administrative hospital data in this way is that its monthly submission allows timely analysis of the data, enabling problems to be dealt with before they can escalate.

Also beneficial when monitoring patient safety is the examination of multiple years of routine data to reveal new trends. Zhang and colleagues (2009) followed the care of elderly patients with an admission for an adverse drug reaction for three subsequent years, to identify factors that increase the risk for further adverse reactions. The study illustrated the fact that for any research conducted, careful interpretation of the data is required due to the periodic update of the classification and coding standards. In addition, the coding rules often differ between states and this must be considered prior to comparing state data (Michel 2008). Interpretation of coded data should be carried out with advice from a trained clinical coder or Health Information Manager to ensure comparability of data over time.

All these patient safety studies involved large samples that were representative of large populations. Studies of this size are only feasible using routine data; medical record review studies are generally much smaller as the process is lengthy and costly. Medical records are also not as accessible to researchers as de-identified routine data due to the ethical and confidentiality issues that have to be addressed prior to access. However, collected routine data is only useful if it is accurate and complete, and therefore maintaining a high standard of clinical documentation and clinical coding is important.

**Risk adjustment**

Australian routine data have been used to develop and test risk adjustment algorithms which support patient safety and other clinical research. Risk adjustment considers a patient’s comorbidities, amongst other variables, to estimate the probability of an adverse outcome. Once developed, these algorithms attempt to distinguish comorbidities from complications in the routine data and allow a patient’s outcome to be more accurately predicted.

The Charlson index (Charlson et al. 1987) was developed using ICD-9-CM codes to measure the severity of a patient’s condition based on their comorbidities, in order to predict the probability of in-hospital mortality. The Index was subsequently modified by Deyo (Deyo, Cherkin & Ciol 1992). The Deyo coding algorithm, which includes codes representing common comorbidities, was later translated into ICD-10-AM by Australian researchers (Sundararajan et al. 2004). Victorian hospital data was used to validate the new algorithm, and it has been used by Ehsani (Ehsani, Jackson & Duckett 2006; Ehsani, Duckett & Jackson 2007) to avoid confounding costs attributable to a patient’s comorbidities with those attributable to complications of care. The Charlson Index and its derivatives have also been
used overseas in clinical research to allow more accurate prognosis of disease.

In an effort to improve on the Charlson Index, Holman et al. (2005) used Western Australia’s linked patient admissions from 1989-1997 to conduct a longitudinal study to develop a more comprehensive comorbidity scoring system. Comorbidities were included in the new Multipurpose Australian Comorbidity Scoring System (MACSS) if they were found to contribute to outcomes such as one-year mortality, 30-day readmission and longer than expected length of stay (Holman et al. 2005). The performance of the MACSS was tested against the Charlson Index using specific subgroups of patients (i.e. Asthma, acute myocardial infarction [AMI], transurethral resection of the prostate [TURP]), from the Western Australian administrative data set.

For the Queensland Health VLAD indicators reviewed above, specific sets of comorbidities were selected to risk-adjust each VLAD Indicator based on a number of criteria, including the frequency of incidence within that patient cohort (Clinical Practice Improvement Centre of Queensland Health 2008). This information was determined using the Queensland hospital admitted patient data. Another Queensland study tested enhancing the administrative data with clinical variables to calculate fatality rates after admission for AMI (Johnston et al. 2007). The collection of additional clinical items was considered to be a potentially valuable supplement, but was concluded to be expensive in comparison with the use of the administrative database alone, which already includes some of the clinical elements sought (Johnston et al. 2007).

The studies reviewed here relied on the recording of all patient comorbidities in the medical record, if meeting criteria for the additional diagnosis standard, to be coded accurately and completely, regardless of whether or not the inclusion affects DRG allocation. The introduction of ‘condition onset’ flagging is set to improve risk-adjustment algorithms further, by restricting the inclusion of diagnoses to those flagged as ‘present on admission’. In the past, diagnosis codes representing conditions, such as anaemia and urinary tract infections, would have been considered as ‘comorbidities’ regardless of the timing of onset, as there was no method of distinguishing conditions present on admission from those arising during the admission.

**Injury surveillance**


Unfortunately the external cause data has not been of optimum use to injury researchers in the past due to the lack of specificity of the codes assigned (McKenzie et al. 2006). In part this can be attributed to substandard coding, but incomplete documentation is a well known contributor to the problem. In a recent survey involving clinical coders, missing information was cited as the major reason for the lack of specificity (McKenzie et al. 2008). The addition of more specific codes to ICD-10-AM was recommended to increase specificity for the purpose of injury surveillance. Broad categories are one of the drawbacks of classification systems such as ICD-10-AM.

Administrative data, despite limitations, are still used in trauma and injury research to evaluate public health policies and to prioritise resources. However, in order to improve injury surveillance in Australia, data recommendations listed in the National Injury Prevention and Safety Promotion Plan 2004-2014 would need to be implemented (Mitchell et al. 2008).

**Incidence and burden of disease**

Incidence of disease can be difficult to determine from admitted patient data as only those patients admitted for treatment can be identified (Hargreaves 2001). Patients with illnesses such as colorectal cancer and sepsis, are likely to be admitted to hospital due to the serious nature of their condition. In these instances, incidence can be estimated from hospital morbidity data. However, cases of death prior to admission are
Reports

excluded and this will result in underestimation of the incidence rate.

A Western Australian study has used hospital morbidity data linked to the cancer registry and mortality records to determine the incidence, mortality and outcomes of colorectal cancer patients in the state (Semmens et al. 2000). A fourteen-year collection of data was used to establish that there had been an increase in incidence of rectal cancer in men and an increase in the mortality of colon cancer in women during that time.

Analysis of four years of routine morbidity data was used to determine the characteristics and incidence of sepsis in Victoria. The results revealed similarities with published North American and European rates, indicating a high degree of reliability in the Australian data (Sundararajan et al. 2005).

Epidemiological studies such as these are important in identifying the extent and burden of certain diseases in Australia. It is difficult to determine incidence without linked hospital data, especially if the disease is chronic, as patients may be double counted with multiple admissions (Brameld et al. 2003). A unique patient identifier used across facilities (or probabilistic record linkage as in Western Australia) would advance the use of administrative data in this role.

Clinical outcomes research
Holman and colleagues (2008) have recently reviewed the creation of and research arising from the Western Australian Data Linkage System (WADLS) since 1995. Studies on treatment outcomes for colorectal (Semmens et al. 2000) ovarian (Laurvick et al. 2003), breast (Hall & Holman 2003), and prostate (Hall et al. 2005) cancers; cholelithiasis (Fletcher et al. 1999), urolithiasis (Holman et al. 2002) and bariatric surgery (Smith et al. 2008) are examples of clinical research relying in part on the use of coded inpatient data. In these studies, cohorts of patients are usually identified from index admissions for particular diagnoses, with patient care followed through the linked data for a specified period.

Health economics research
Routinely-coded hospital data can be combined with information on individual patient treatment costs available from sophisticated computerised costing systems. Not all hospitals have invested in ‘clinical costing’ or ‘activity-based costing’ technologies, but where they have, the information can be used in a range of health economics studies. Carlin, for example, used cases with a primary diagnosis of gastro-enteritis at a Melbourne children’s hospital to estimate hospitalisation costs that could be saved by introduction of rotavirus vaccination (Carlin et al. 1999).

Costs of inpatient adverse events have also been estimated using the routine data. Ehsani and colleagues (2006) identified cases with at least one C-prefixed diagnosis in Victorian data to estimate that a hospital-acquired diagnosis adds nearly $7,000 to the episode costs, after taking into account the patient’s DRG and Charlson Index co-morbidities. Related work by Ehsani et al. (2007) focused on complications in cardiac surgery episodes, selecting patients on the basis of their admitting diagnosis and comparing those with and without a C-prefix. For patients with multi-day stays, complications were found to add seven days to the patient’s stay, and add nearly 30% to episode costs.

Western Australian research by Calver et al. (2006) used data linkage capacities in that state to identify ‘high cost’ users of inpatient care. Analysing principal diagnosis information over multiple admissions, researchers found that ongoing care, including dialysis, chemotherapy, angina and heart failure resulted in the highest annualised costs of care. High-cost users represented only 26% of total admissions, but 38% of inpatient costs.

Conclusion
The aim of this article was to demonstrate how important administrative data, and hence the work of clinical coders, is to health care in Australia. Increasingly, the routine inpatient data are being used in Australia to inform clinical management and health policy. Administrative data provides a cost-effective way to monitor large populations on a timely basis.

The quality of the coded data is often questioned (Powell, Lim & Heller 2001; Scott & Ward
2006), based on studies from health care systems where incentive systems are different and less investment is made toward ensuring data quality. Such criticisms are often due to misunderstandings about the strengths and limitations of a standardised and replicable classification system. It has long been argued that variable investment in coding effort across the state-based public hospital systems has led to inconsistencies in data quality in some areas (Michel 2008). Despite this, data audits conducted in Victoria have shown the data to be valid and reliable (Henderson, Shepheard & Sundararajan 2006), and Western Australia and Queensland have taken steps based on auditors’ comments to improve the quality (Stevens et al. 1998; Logan et al. 2006).

Work should continue to safeguard the quality of routine data and to support its increasing use in injury surveillance, clinical outcomes research, economic evaluation and patient safety improvement.

### References


ICD-10-AM/ACHI/ACS  
Sixth Edition

ICD-10-AM/ACHI/ACS Sixth Edition is now available in printed volumes and as eCompress® desktop software.

Select the version that best suits your needs
Printed volumes are available individually or in sets. Printed sets can be purchased with a slipcase (for your bookshelf or desk) or with a durable PVC carry case.

Sixth Edition software
The popular eCompress® version of ICD-10-AM/ACHI/ACS contains all the volumes plus current 10-AM Commandments from Coding Matters. The eBook version allows you to search all volumes, write and share notes, create bookmarks and much more.

Visit the NCCH website for further information and details of how to order your copy.

www.fhs.usyd.edu.au/ncch

---

Health Information Management Journal

SAVE UP TO 20%
on your HIMAA membership renewal

Purchase HIMAA membership 2 or 3 years in advance and receive discounts based on your length of continuous membership/advance purchase.*
Contact Rose Wong on 02 9887 5001 or at himaa@himaa.org.au for more details.

* 6 years continuous membership + 3 years advance purchase means a discount of 20%.

GET A DISCOUNT OF $20
on your HIMAA Conference Registration

Recruit new Full, Associate or Organisational members to HIMAA and receive a discount of $20 per new member off your registration to the next HIMAA national conference/symposium.