Adverse Events Regional Feasibility Study: methodological results

Peter Davis, *Professor, Department of Public Health and General Practice, Christchurch School of Medicine, University of Otago, Christchurch;* **Roy Lay-Yee**, *Analyst, Department of Community Health, School of Medicine, University of Auckland, Auckland;* **Stephan Schug**, *Associate Professor, Department of Anaesthesia, University of Western Australia, Perth;* **Robin Briant**, *Clinical Director, Department of Community Health, School of Medicine;* **Alastair Scott**, *Professor, Department of Statistics;* **Sandra Johnson**, *Project Manager;* **Wendy Bingley**, *Data Manager, Department of Community Health, School of Medicine, University of Auckland, Auckland.*

Abstract

Aims. To assess the feasibility of research into the occurrence, causation and prevention of adverse events (AEs) in New Zealand public hospitals.

Methods. A two-stage retrospective review was carried out on 1575 medical records selected by systematic list sample from admissions for 1995 in three public hospitals in the Auckland region. Following initial screening, medical records were subject to structured implicit review using a standardised protocol. Feasibility measures, using international benchmarks where possible, were: adequacy of sample selection; completeness of medical records; reliability and validity of screener and reviewer judgements; internal consistency and face validity of AE determination and preventability assessment.

Results. The sample selection procedure was effective, although nearly 10% of records could not be secured. Information in medical records was sufficient for the identification and analysis of AEs. Adequate levels of agreement were achieved for screener and reviewer judgements, with kappa scores ranging between 0.302 and 0.622 and positive predictive values between 50.0% and 89.7%. The criteria for AE determination showed internal consistency and face validity, as did those for preventability.

Conclusions. Research into the occurrence, causation and prevention of AEs in New Zealand health care settings is methodologically feasible and meets international benchmark standards.

NZ Med J 2001; 114: 200-2

In New Zealand the question of patient safety has, to date, been the subject of relatively little systematic research. One of the first studies to use a standardised, epidemiological approach was a survey of adverse drug events among over 9000 admissions to Dunedin hospital in the early 1970s.[1] While useful research since that time has been carried out on surgical audit[2] and anaesthetic error,[3] and while the Ministry of Health has published some standardised information across New Zealand hospitals,[4] no generic, epidemiological data on adverse events (AEs) has been published in this country. The absence of such data has been recognised as an obstacle to developing proposals for the regulation of safety in health and disability.[5]

A major scientific stimulus to rigorous epidemiological research on patient safety has been the development of standardised procedures for the assessment of AEs using medical records. It was not until the Harvard Medical Practice Study (HMPS) that a measurably reliable, valid and generic definition of AEs was first established across a wide range of clinical settings.[6] This approach has been replicated in the Quality in Australian Health Care Study (QAHCS).[7]

The object of this study was to test the feasibility of applying a standardised protocol to the analysis of medical records with a view to determining the occurrence, causation and prevention of AEs in New Zealand public hospitals.

Methods

Sampling and data collection. Three major public hospitals were selected for study in the Auckland region. The survey population was defined as all patient admissions to these hospitals for calendar year 1995 (excluding day and psychiatric cases). The sampling frame for each hospital was a list of all eligible admissions. New Zealand Health Information Service (NZHIS) selected a systematic list sample of 525 admissions from each hospital, with cases ordered by admission date. Each selected case signalled an index admission. To be included in the study an AE had to be related to, or occur during, the index admission.

Standard hospital inpatient information for each sampled admission was provided by NZHIS. This included admissions information (dates of admission and discharge, admission type and source), socio-demographic data (age, gender, ethnicity, domicile code), and clinical data (ICD9 and AN-DRG3.1 diagnostic classifications).

The core data collection procedure was a two-stage retrospective review of medical records for selected cases using the Review Form 1 (RF1) and Review Form 2 (RF2), both closely modelled on the comparable instruments in the American and Australian studies.[6,7] The first stage was the RF1 screen undertaken by Registered Nurses (RNs). The purpose of this stage was to ascertain if the hospitalisation in question - the index admission - met any of eighteen screening criteria selected as potentially indicative of an AE.[8] The second stage undertaken by Medical Officers (MOs) used the RF2, an instrument relying on structured implicit review (that is, the guided exercise of professional judgement), and was designed to determine the presence and context of any AE. In two of the three hospitals an Expert Reviewer (ER) administered "blind" the full cycle of data collection on a one-in-ten sub-sample.

Definitions.[8] An AE was operationally defined as (a) an unintended injury or unintended complication, (b) resulting in temporary or permanent disability, including increased length of stay and/or financial loss to the patient, (c) that was caused by health care management rather than the underlying disease process. A key part of AE determination was the assessment of the extent to which an identified injury resulting in disability was caused by health care management. In order to assist reviewers to make this judgement they were guided through a series of seven evaluation questions.

Preventability of an AE was assessed as an error in health care management due to failure to follow accepted practice at an individual or system level. In order to assist reviewers in making a judgement about the preventability of AEs they were required to work through ten evaluation questions.

Evaluation of feasibility. (1) Adequacy of sampling – comprehensive sample frame and "success rate" in accessing records. (2) Completeness of records – information available for data collection. (3) Reliability – kappa and positive and negative predictive values (MO as criterion). (4) Validity – kappa and positive and negative predictive values (ER as criterion). (5) Internal consistency and face validity – assessment of AE determination and level of preventability against individual items using positive and negative predictive values (analysis of AE status against discharge mode, ICD external cause code and length of stay, was also carried out, but these results are not reported in full).

Where possible, benchmark comparisons will be made with the corresponding data drawn from QAHCS.[8]

Results

Adequacy of sampling. Of the 246 sampled records that could not be screened, 30.1% could not be retrieved, 24.4% had inadequate documentation, and 45.5%, mainly day stay, were incorrectly included in the sample by NZHIS. There

were also three records screened criteria positive but not available for further review. Excluding the mis-sampled admissions, the success rate was 90.8%.

Completeness of medical records (Table 1). For the first stage of the review procedure, the RN screen, the available information was judged to be sufficient to complete all aspects of the RF1 in nearly 95% of all sampled records. For the second stage, the MO review, the available information was deemed sufficient to complete all aspects of the RF2 for nearly 85% of all cases classed as AEs, and in the remainder was adequate to determine AE occurrence.

Table 1. Adequacy of medical records.

Medical Record Items	RN* Screening (RF1†): Percentage of screened admissions	MO‡ Reviewing (RF2§): Percentage of AE admissions
Initial medical assessment	99.5 % (n=1326)	96.5 % (n=142)
If applicable:		
Medical progress notes	98.7%	91.5%
Nursing progress notes	99.4%	97.9%
Procedure documentation	99.2%	97.2%
Pathology reports	98.6%	98.6%
Discharge summary	96.3%	94.4%
All above items adequate	94.1%	84.5%

* Registered Nurse. † Review Form 1. ‡ Medical Officer. § Review Form 2.

Reliability and Validity (Table 2). Reliability showed only moderate results. While agreement on criteria presence was high (89.7%), the positive predictive value for AE presence was a little over 50%. There was a similar pattern with validity. There was a low level of agreement for screeners on AE presence. However, positive predictive value and kappa scores were acceptable for criteria presence and AE determination. The small number of cases for AE determination should be noted.

Table 2. Reliability and validity of screener and reviewer judgements.

	Kappa	Percent agreement	Positive predictive value	Negative predictive value
Reliability				
RN*/MO†: criteria presence‡ (n=553)	-	89.7%	89.7%	-
RN/MO: AE presence (n=548)	0.344	74.8%	51.4%	83.0%
Validity				
RN/ER§: criteria presence (n=74)	0.465	74.3%	71.4%	76.1%
RN/ER: AE presence (n=72)	0.302	86.1%	50.0%	89.4%
MO/ER: AE determination (n=28)	0.622	85.7%	62.5%	95.0%

* Registered Nurse. † Medical Officer. ‡ Only cases screened positive by RNs were further reviewed by MOs; thus kappa and negative predictive value are not applicable. §Expert Reviewer.

AE Determination (Table 3 and Figure 1). The results of reviewer responses on AE determination are presented. The first question, whether there was a note in the record suggestive of the causal role of health care management, was strongly predictive of an AE (positive predictive value=92.7%). In the case of the second question – a note predictive of injury – the relationship was weak. The assessment of the timing of events was the only other item that was strongly predictive of a reviewer's attribution of an AE. The remaining questions showed a moderate tendency to be predictive of an AE.

Following these seven evaluation questions reviewers were then required to make an assessment of the degree to which the outcome was 'caused' by health care management. The results of this exercise are outlined in Figure 1. For nearly half of all cases with both injury and disability or longer hospital stay there was virtually no evidence of health care management causation in the opinion of the reviewer. These were excluded from further analysis. The full protocol - RF2 - was administered to the remaining 142 cases. It is notable, however, that only 70 of these showed moderate, strong or virtually certain evidence of health care management causation.

Table 3. Adverse event status by evaluation category: percent agreement and predictive value.

Evaluation Category [POSITIVE/negative]	Adverse Event Status (n=235) [Present (n=142)/ Absent] Percent agreement	Positive predictive value	Negative predictive value
*Q. Is there a **note** in the medical record which indicates or **suggests** that **health care management** caused the injury?* YES/no	79.6%	92.7%	67.5%
*Q. Is there a **note** in the medical record which **predicts** the possibility of an **injury** from the patient's disease?* NO/yes	56.5%	65.2%	44.0%
*Q. Does the **timing** of events **suggest** that the injury was related to the **treatment**?* LIKELY/possible, unlikely	76.8%	83.6%	67.7%
*Q. Are there **other reasonable explanations** for the cause of the injury?* FEW/some, many	63.5%	75.9%	51.8%
*Q. Was there an **opportunity** prior to the occurrence of the injury for **intervention** which might have **prevented** it?* YES/possibly,no	52.0%	75.9%	43.6%
*Q. Is there **recognition** that the **intervention** in question causes this kind of **injury**?* WIDELY/recognised by other specialists,no	62.3%	81.1%	32.1%
*Q. Did the adverse event **respond to new management** to neutralise or modify the effects of former management?* CONVINCING/suggestive,no	60.5%	75.2%	37.3%

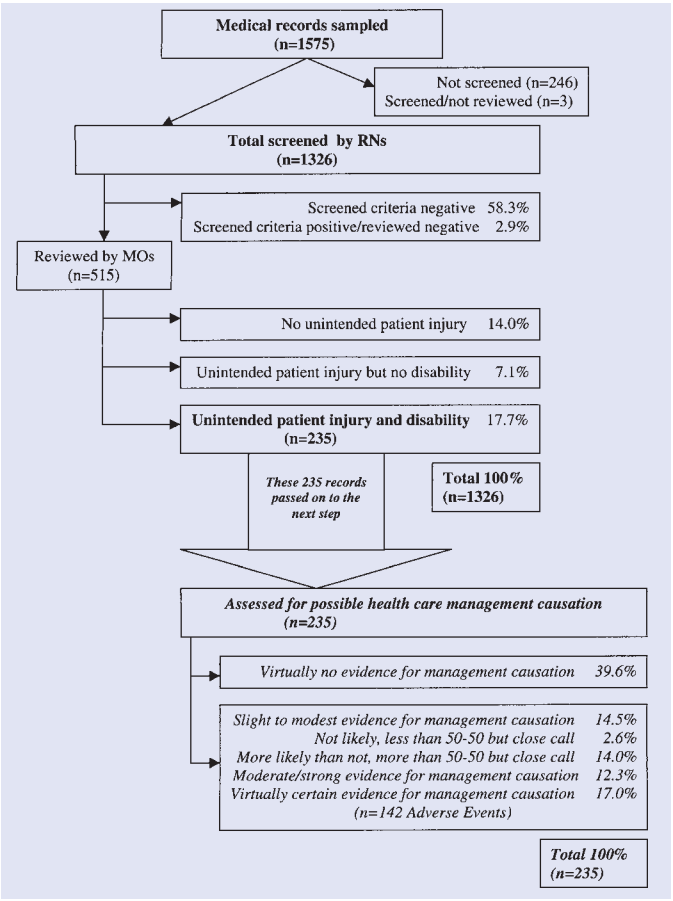

Figure 1. Assessing health care management causation.

Assessment of preventability (Table 4). The results of reviewer responses on assessment of preventability are presented.

Whether there was consensus about diagnosis and therapy had little bearing on their judgement of preventability (positive predictive value=46.0%). Complexity, co-morbidity, degree of emergency, potential benefit, chance of benefit, and risk of an AE were other questions with little predictive value. By contrast, appropriateness of management, deviation of management from the accepted norm, and reflection on repetition were questions that were more predictive of reviewers' judgements of high preventability.

Table 4. – High preventability* of adverse events by evaluation category: percent agreement and predictive value.

Evaluation Category [POSITIVE/negative]	High Preventability of Adverse Events (n=142) [Present (n=55)/ Absent]		
	Percent agreement	Positive predictive value	Negative predictive value
*Q. Is there **consensus** about diagnosis and therapy regarding this case?* GREAT DEAL/some, very little	56.4%	46.0%	73.6%
*Q. How **complex** was the case?* UNCOMPLICATED/moderate, very	58.2%	45.8%	64.5%
*Q. Was the **management** in question **appropriate?*** NOT, POSSIBLY/probably, definitely	75.0%	75.7%	74.8%
*Q. What was the **co-morbidity** of the case in which the adverse event occurred?* NONE/moderate, very ill	51.8%	38.2%	60.5%
*Q. What was the degree of **deviation of management** from the accepted norm?* SEVERE, MODERATE/little	69.8%	67.6%	70.6%
*Q. What was the degree of **emergency** in management of the case prior to the occurrence of the adverse event?* NONE/moderate, critical	58.6%	46.7%	64.2%
*Q. What **potential benefit** was associated with the management?* MINOR/major, life-saving	52.8%	23.8%	58.5%
*Q. What was the **chance of benefit** associated with the management?* HIGH/moderate, low	51.2%	38.8%	65.0%
*Q. What was the **risk of an adverse event** related to the management?* HIGH, MODERATE/low	50.8%	41.3%	63.2%
*Q. On reflection, would a reasonable doctor or health professional **do this again**?* NO,PROBABLY NOT/ probably,definitely	72.5%	72.7%	72.4%

* Preventability judged to be more likely than not.

Discussion

The primary objective of this study was to assess the feasibility of conducting research into the occurrence, causation and prevention of AEs .

In methodological terms the feasibility study was able, in the first instance, to establish the adequacy of the sample frame and the effectiveness of the sampling procedure. In QAHCS,[7] for example, the sample frame had to be constructed for each hospital. Through NZHIS we were able to draw samples centrally. The success rate - that is, the proportion of sampled records which was screened – was, however, relatively low; 90.8% compared with the Australian rate of 96.8%.[8]

The level of the information available appeared to be adequate, and comparable to international results. Thus, the standards of medical documentation were sufficient to permit almost universal completion of the RF1. In the case of the RF2, while the determination of AE status was possible in virtually all cases, at least one data item was missing in a sixth of completed RF2s. These results are comparable to those achieved in the Australian study.[8] The quality of the assessment process – that is, the level of agreement on the screening and reviewing tasks - was adequate when compared to the results from QAHCS.

More important to the validity of the study was the process of AE determination. In QAHCS and HMPS two MO reviewers were used, with arbitration in case of disagreement. Despite this, the overall level of agreement in these studies was low.[9] This confirms the conclusion of one authority that physician agreement on the quality of care is often only slightly better than chance.[10] Furthermore, research suggests that discussion between reviewers does not actually improve the reliability of peer review of hospital quality.[11] In this investigation, as in the more recent Utah and Colorado studies (UTCOS), only a single MO was used, with 'blind' expert review of a 10% sub-sample. The positive predictive value for agreement between MO reviewers and ERs - admittedly on a small sample - was 62.5% (kappa=0.622). This equates to the level of agreement achieved between MOs in QAHCS (kappa=0.55)[8] and between reviewers in a reliability study of a sub-sample in UTCOS (kappa=0.4).[12]

The study broke relatively new ground in attempting to go beyond conventional measures of internal validity. Thus it was possible to use the evaluation items in the assessment of AE status and of preventability to establish the internal consistency and face validity of reviewer judgements. For example, of the records assessed for AE status and which had a note indicating that health care management was causative, over 90% were judged by reviewers to be AEs. Similarly, where the reviewer deemed management inappropriate, a high level of preventability was a likely assessment.

In a separate exercise not reported here, internal consistency was also assessed against routinely-collected hospital information. In this case, however, results did not prove to be useful. In particular, external cause codes signalling medical misadventure had low sensitivity, identifying only a quarter of AEs, although length of stay was more predictive (a greater than average stay was associated with half all AEs).

In conclusion, this study was designed to test at a regional level the feasibility of carrying out research on the detection and analysis of AEs. The investigation has demonstrated sufficient levels of performance in methodological terms, as judged by international benchmark standards for work of this kind.

Acknowledgements. This study was funded by the Health Research Council of New Zealand. We are very grateful to the management and clinical staff of the three participating hospitals and to the study's Advisory and Monitoring Committee, chaired by Dr. David Richmond. Specifically, we thank Dr Colin McArthur of Auckland Hospital, Dr Ian Brown and Dr Peter Gow of Middlemore Hospital, and Dr Andrew Love of North Shore Hospital, for their assistance and comments. Valuable comments were also contributed by an anonymous reviewer. We wish to thank our medical review and data processing teams for their meticulous work, and hospital records staff for their willing co-operation. Finally, we acknowledge the assistance of statistics students supervised by Professor Scott.

Correspondence. Professor Peter Davis, Department of Public Health and General Practice, Christchurch School of Medicine, University of Otago, PO Box 4345, Christchurch, New Zealand. Fax: (03 3640 425; email: Peter.Davis@chmeds.ac.nz

1. Smidt NA, McQueen EG. Adverse reactions to drugs: a comprehensive hospital inpatient survey. NZ Med J 1972; 76: 397-401.
2. Pettigrew RA, McDonald JR, van Rij AM. Developing a system for surgical audit. Aust NZ J Surg 1991; 61: 563-9.
3. Merry AF, Peck DJ. Anaesthetists, errors in drug administration and the law. NZ Med J 1995. 108: 185-7.
4. Ministry of Health, 1995, Purchasing for Your Health 1994/95. Wellington; 1995.
5. Ministry of Health. Taking care II. Wellington; 1997.
6. Harvard Medical Practice Study. 1990, Patients, doctors and lawyers: medical injury, malpractice litigation, and patient compensation in New York. New York: Harvard College; 1990.
7. Harvey R, Goss J. Report on the feasibility study of an Australian hospitals' adverse health care incidents study. Canberra, Australia: Commonwealth Department of Health, Housing and Community Services; 1992.
8. Wilson RM, Runciman WB, Gibberd RW et al. The quality in Australian health care study. Med J Aust 1995; 163: 458-71.
9. Localio AR, Weaver SL, Landis JR et al. Identifying adverse events caused by medical care: Degree of physician agreement in a retrospective chart review. Ann Intern Med 1996; 125: 456-64.
10. Goldman RL. The reliability of peer assessments of quality of care. JAMA 1992; 267: 958-60.
11. Hofer TP, Bernstein SJ, DeMonner S, Hayward RA. Discussion between reviewers does not improve reliability of peer review of hospital quality. Med Care 2000; 38: 152-61.
12. Thomas EJ, Studdert DM, Burstin JR et al. Incidence and types of adverse events and negligent care in Utah and Colorado. Med Care 2000; 38: 261-71.