

Effectiveness of routine reporting to identify minor and serious adverse outcomes in surgical patients

P J Marang-van de Mheen, N van Hanegem, J Kievit

Qual Saf Health Care 2005;14:378–382. doi: 10.1136/qshc.2004.013250

See end of article for authors' affiliations

Correspondence to:
Dr P J Marang-van de Mheen, Department of Medical Decision Making, Leiden University Medical Centre, J10-S, PO Box 9600, 2300 RC Leiden, The Netherlands; p.j.marang@lumc.nl

Accepted for publication
25 June 2005

Objective: To assess the effectiveness of routine reporting to identify surgical adverse outcomes in comparison with retrospective medical record review.

Design: Independent assessment of two methods applied to one sample. Surgeons and surgical residents routinely reported all adverse outcomes for patients in their care during admission. A trained research assistant, blinded to the surgeons' reporting data, retrospectively reviewed the medical records of selected patients and registered all adverse outcomes identified from paper or electronic patient records.

Setting: Dutch university hospital.

Study sample: A 5% sample of patients (N = 150) discharged in 2002 was taken; oversampling of patients undergoing reoperations, sick patients (ASA \geq 3), and those undergoing technically complex surgery was done to increase the yield of adverse outcomes.

Main outcome measures: The number of adverse outcomes identified by each method was compared with the total number identified by either method. This was done both for all adverse outcomes and for serious adverse outcomes.

Results: Routine reporting identified fewer adverse outcomes than medical record review (62.5% v 78.2%). Complete agreement was achieved in only 40.7% of adverse outcomes. Routine reporting identified slightly more serious adverse outcomes (84.8% v 79.5% of the total), but this difference was not statistically significant. Extrapolating these results to the total number of admissions in 2002, routine reporting underestimated the annual adverse outcome incidence by 1.8% (increasing from 14.5% to 16.3%) and the incidence of serious adverse outcomes by 0.3% (increasing from 6.9% to 7.2%).

Conclusions: Neither method identified all adverse outcomes. Routine reporting underestimated the incidence of minor events but was as accurate as record review in identifying serious adverse outcomes.

Most studies on adverse outcomes rely on retrospective medical record review to identify events and outcomes.^{1–5} The rationale for using record review is that medical records are thought to be complete and therefore provide the best possible information. However, record review is time consuming and may be distorted by missing written information. In addition, underlying reasons for choices and actions each associated with a certain risk may remain unknown and lead to inter-observer variation in judgements on the preventability of adverse outcomes.⁶

In the Netherlands nationwide routine reporting of surgical adverse outcomes integrated in daily medical care was started in 1998.⁷ As reporting is part of daily care, all the necessary information is available at the time of actual registration and encoding of adverse outcomes. In routine reporting all adverse outcomes are identified by surgeons and surgical residents whereas the medical records may also contain information reported by nurses or other caregivers. In addition to routine reporting, there is a weekly meeting encompassing a causal analysis aimed at identifying methods of prevention. Because of this weekly meeting and the fact that reporting is carried out by surgeons themselves as part of daily practice, it may have a greater potential for improving the quality of care than a retrospective review by an outsider. However, busy doctors may forget to report adverse outcomes—even serious ones—making them not the most suitable data managers (Veltkamp SC, PhD thesis, unpublished, 2001).

The purpose of the present study was to compare the effectiveness of routine reporting and medical record review in identifying minor and serious adverse outcomes in surgical patients. The study was carried out in a sample of 150

patients admitted to a university hospital in which high risk patients were oversampled to increase the yield of adverse outcomes. This is valid since the purpose is to compare the two methods, but the results therefore cannot be interpreted as a true incidence of adverse outcomes. To assess the impact of these results on the overall adverse outcome incidence, the results of the present study were extrapolated to the entire discharged population.

METHODS

Definitions

In the Netherlands medical specialists have agreed on one common definition of adverse outcomes.⁸ This definition differs from that used in other studies because it has been chosen with the explicit aim of excluding subjective judgement on cause and effect, and right and wrong. The definition of an adverse outcome is:

“... an unintended and unwanted event or state occurring during or following medical care, that is so harmful to a patient's health that (adjustment of) treatment is required or that permanent damage results. The adverse outcome may be noted during treatment or in a predefined period after discharge or transferral to another department. The intended result of treatment, the likelihood of the adverse outcome occurring, and the presence or absence of a medical error causing it are irrelevant in identifying an adverse outcome.”

This definition is more sensitive than the definitions used in other studies in which adverse outcomes linked to the natural history of the disease as well as co-morbidity are usually excluded.^{1–4} The advantage of the above definition is that it is less subject to discussions on whether or not this could be expected given the medical condition of the patient, and therefore less likely to result in inter-observer variability.

Conditions already present at admission are excluded by this definition and are thus not considered adverse outcomes. Surgeons routinely report adverse outcomes occurring in patients admitted under responsibility of the Department of Surgery and/or operated by a surgeon throughout the admission period up to 30 days after discharge. Reported adverse outcomes are entered into dedicated software and are automatically encoded on a three-dimensional encoding system according to (1) nature (e.g. infection), (2) location (both topographically, e.g. thorax, and with respect to organ or organ system, e.g. respiratory tract), and (3) determinants and other relevant information (e.g. medication). Furthermore, for each adverse outcome the severity of its impact on the patient's health is registered in four categories: (1) temporary health disadvantage, recovery without (re)operation; (2) recovery after (re)operation; (3) (probably) permanent damage or function loss; and (4) death.

In the present study category 1 was defined as minor adverse outcomes and categories 2, 3 and 4 as serious adverse outcomes.

Collection of data

Routine reporting data were collected in the Department of Surgery at the Leiden University Medical Centre where all surgical specialties (such as vascular surgery and oncological surgery) are represented. There is a long tradition of adverse outcome reporting in this department going back to the beginning of the 1980s, which means that all clinically occurring adverse outcomes consistent with the above definition are systematically reported on special forms in the patients' records. The presence and completeness of each patient's form is checked weekly. Written information on adverse outcomes is checked and discussed 2 weeks after discharge at a weekly meeting attended by all surgeons, residents, interns, and medical students. The purpose of the discussion is twofold: (1) to assess whether the written information is understandable and correct; and (2) to analyse, interpret and conclude whether anything can be learned from this particular case. After consensus has been reached on the correct information, adverse outcomes are entered into the dedicated software. The database averages about 2750 patient admissions per year.

From the 2899 patients discharged in 2002, a random sample of 150 patient admissions (5.2% of the total population of that year) was drawn, selecting no more than one admission in patients with repeated admissions. In order to increase the yield of adverse outcomes we oversampled the following patients:

- patients with a reoperation within the same admission since these may be treatment of adverse outcomes;
- high risk patients at the start of admission, defined as patients with an ASA class⁹ of 3 or higher at the first operation of admission; and
- patients undergoing technically complex procedures defined as a complexity class of 5 or higher on a scale from 1 (the technically relatively simple procedures) to 7 (technically very complex or difficult).¹⁰

To achieve this oversampling we selected 52 from the 210 patients (24.8%) undergoing a reoperation during the same admission. A relatively high percentage was sampled from this group because (1) it is likely that these reoperations occur to treat adverse outcomes (and the aim was to increase the yield of adverse outcomes) and (2) the group was rather small so we had to sample a quarter of these patients to obtain a representative sample from this group (rather than selecting patients with planned reoperations). From the 816 high risk patients or patients undergoing complex

procedures, 57 (7%) were selected. This percentage is smaller because of the greater number of patients in this group. From the remaining 1873 patients, 41 were selected (2.2%) to arrive at the total of 150 patients.

A specially trained research assistant with medical training reviewed the medical records of all the selected patient admissions, including both paper records and data from the hospital information system, and registered all events and outcomes that fitted the definition. The research assistant was blinded to the surgeons' reporting data and to the sampling information. The independence of the two methods was ensured by removing the surgeons' reporting forms from the patients' charts at the time of the retrospective record review. If no information in the surgical charts was found regarding the selected admission, we looked for information in the charts from the outpatient department or other specialties. Consistency of the record review data with the definition was checked and any inconsistencies were discussed between the research assistant, an epidemiologist, and a surgeon until consensus was reached. The medical records concerning the selected admission of one high risk patient could not be retrieved, thereby leaving 149 patient admissions. All medical record review information was written on a predesigned form and entered into a separate database.

Analysis of data

The effectiveness of the two methods was calculated as the number of adverse outcomes identified by each method divided by the total number of adverse outcomes identified by either method. This was done both for all adverse outcomes and for serious adverse outcomes. Since adverse outcomes cluster within admissions, additional analyses were performed for the number of admissions with and without adverse outcomes. χ^2 tests were used to test for differences in the effectiveness of the two methods.

The results were extrapolated to the total discharged population in 2002 to show how the results of the present study would affect the adverse outcome incidence as reported by surgeons. Given the oversampling of high risk and complex patients and the fact that adverse outcome probabilities are higher in these groups, the impact on the overall incidence may be much smaller than estimated in the sample.

RESULTS

Table 1 shows the basic characteristics of the patient population discharged in 2002 and of the sample used in the present study. Sample patients were older, were more often operated, had a higher risk at the start of the first operation, and underwent more complex surgery, all differences caused by the intentional oversampling.

A total of 381 adverse outcomes was identified by either method, 238 (62.5%) by routine reporting and 298 (78.2%) by record review; complete agreement was achieved in 155 (40.7%, table 2). Record review identified significantly more adverse outcomes overall than routine reporting (paired $\chi^2 = 15.93$, $p < 0.001$). Of 112 serious adverse outcomes, routine reporting identified 95 (84.8%), record review 89 (79.5%), and full agreement existed in 72 (64.3%, table 2). Although routine reporting identified more severe adverse outcomes, this difference did not reach statistical significance (paired $\chi^2 = 0.9$, $p = 0.34$).

With respect to admissions with or without adverse outcomes, there was agreement between both methods that adverse outcomes occurred in 48 admissions (although they might identify different numbers of adverse outcomes within an admission), and that 76 admissions were event-free (table 3, kappa = 0.66). Both methods were discrepant on 25

Table 1 Basic characteristics of patient population: total population discharged in 2002 versus sample in present study

	Total population	Present study	Test of difference	
			Statistic	p value
Number of admissions	2932	149		
Mean (SD) age at admission (years)	48.0 (22.6)	52.7 (21.5)	$t=2.48$	0.013
% male	55.1	57.7	$\chi^2=0.39$	0.53
% admissions with operation	68.3	94.6	$\chi^2=46.33$	<0.001
% with ASA class of first operation ≥ 3	20.7	35.5	$\chi^2=18.45$	<0.001
% with maximum complexity class of first operation ≥ 5	32.9	46.5	$\chi^2=11.2$	<0.001

admissions: in 22 admissions adverse outcomes were only identified by record review and in three admissions only by routine reporting. As a result, routine reporting estimated that adverse outcomes occurred in 34.2% of admissions in this selected patient set compared with an estimate of 47.0% by record review. With respect to serious adverse outcomes only, the figures were 27.5% for routine reporting and 30.9% for record review.

In the total population discharged in 2002, adverse outcomes were reported in 14.5% of all patient admissions. For each category of patients that was used to oversample high risk patients (such as those who had undergone a reoperation), we applied the net underreporting of admissions with adverse outcomes as found in the present study to the reported incidence in the total discharged population. For example, the underreporting of 13.5% in the sample of reoperated patients was used to increase the reported number of 134 admissions with adverse outcomes in the total population of reoperated patients, resulting in an adjusted number of 152 admissions with adverse outcomes. This was also done for the group of high risk patients (16.8% underreporting) and the remaining patients (9.8% underreporting). Adding the adjusted numbers in these groups, we estimated the "true" number of admissions with adverse outcomes to be 472, which is 16.3% of the total number of 2899 admissions. Likewise, the annual incidence of admissions with severe adverse outcomes would increase from 6.9% to 7.2%.

DISCUSSION

The present study has shown that no gold standard exists for identifying adverse outcomes, as adverse outcomes were

missed by both routine reporting and record review. Routine reporting identified fewer adverse outcomes than medical record review (62.5% v 78.2%) and complete agreement existed in only 40.7% of adverse outcomes. However, routine reporting identified slightly more serious adverse outcomes but this difference did not reach statistical significance. Medical record review therefore identified more minor adverse outcomes causing a temporary health disadvantage without a (re)operation being necessary. Extrapolating these results to the total discharged population, we estimated that the incidence of adverse outcomes would increase by 1.8% (from 14.5% to 16.3%), and that of serious adverse outcomes by 0.3% (from 6.9% to 7.2%).

The percentage of identified serious adverse outcomes did not differ between the two methods. This may have been caused by lack of power, given the number of serious adverse outcomes and the small difference between the methods. However, routine reporting identified more serious adverse outcomes than the record review (95 v 89), rather than underreporting as was found for the total number of adverse outcomes. It therefore seems unlikely that increasing the power would completely reverse these results to underreporting of serious adverse outcomes by routine reporting. We conclude that any difference between the two methods that may exist in either direction is too small to be detected.

Other studies do not give unequivocal results. In an unpublished study by Veltkamp, adverse outcomes were reported using four different methods: retrospectively at discharge, by doctors, by nurses, and by a fulltime researcher who checked all the patients daily. Of the 955 adverse outcomes occurring during admission, 299 (31%) were reported by the researcher only and would otherwise have

Table 2 Number of adverse outcomes identified by routine reporting and in medical records according to level of severity

Severity of adverse outcome	Number of adverse outcomes			Total
	Reporting only (%)	Medical records only (%)	Both sources (%)	
Minor adverse outcomes (1) Temporary health disadvantage, recovery without (re)operation	60 (22.3%)	126 (46.8%)	83 (30.9%)	269
Serious adverse outcomes (2) Recovery after (re)operation	19 (20.7%)	14 (15.2%)	59 (64.1%)	92
(3) (Probably) permanent damage or function loss	3 (37.5%)	2 (25.0%)	3 (37.5%)	8
(4) Death	–	1 (9.1%)*	10 (90.9%)	11
Total (%)	83 (21.8%)	143 (37.5%)	155 (40.7%)	381†

*Occurring after transferral.

†Of one adverse outcome identified by routine reporting only, the severity was unknown because of the premature death of the patient due to another cause.

Table 3 Admissions with and without adverse outcomes: routine reporting versus medical record review

Medical record review	Routine reporting		Total
	No adverse outcomes	Adverse outcomes	
No adverse outcomes	76	3	79
Adverse outcomes	22	48	70
Total	98	51	149

been missed; 128 (43%) of these were serious outcomes. In our study 143 (38%) of the adverse outcomes were missed by routine reporting, 17 (12%) of which were serious adverse outcomes. Although these studies differ in design, it is clear that significantly fewer serious adverse outcomes were missed in our study than in the study by Veltkamp, which suggests that underreporting is a hospital- and/or context-dependent phenomenon.

O'Neill *et al*¹¹ studied adverse medical events using a narrower definition focused on medical injuries. They compared a physician reporting system with retrospective record review to identify the adverse medical events. Both methods uncovered equal numbers of adverse events, but with even smaller complete agreement (30.8% uncovered by both methods) than in the present study. The proportion of adverse events found by both methods was even smaller in the study by Wolff *et al*¹² which was attributed to the fact that most events were reported by nursing staff and only a few events by medical staff. One reason why the present study had a higher level of agreement might be that there is an open non-punitive atmosphere for reporting adverse outcomes in our department and doctors are encouraged to report minor adverse outcomes also.¹³ On the other hand, O'Neill *et al* found that the physician reporting system reported more preventable adverse events and would thus be more appropriate for quality improvement.¹¹ Similar results were found by Michel *et al*¹⁴ who compared the effectiveness of a cross sectional, a prospective, and a retrospective method to identify adverse events. A distinction by level of severity has not been included in previous studies.

In the present study agreement between the two methods was only 31% for minor adverse outcomes. As expected, it was higher for serious adverse outcomes (64%). Part of the explanation for the discrepancy between the two methods is that adverse outcomes occurring after discharge or transferral are better noted in the medical records and are not picked up by routine reporting. This explained two of the 17 serious adverse outcomes missed by routine reporting, including one death. Furthermore, most of the missed adverse outcomes were those from which the patient recovered after (re)operation. To improve this, the weekly check of registration forms will include a check on whether reoperations were planned or were the result of adverse outcomes.

Surprisingly, the present study shows that routine reporting may identify (serious) adverse outcomes that are missed by medical record review. This may be because the written records are not complete or because the adverse outcomes were not picked up in the review. We therefore double checked all these adverse outcomes and found that no information was present in the medical records, suggesting incomplete information in the written records. This may be explained by the fact that surgeons report adverse outcomes on the registration form but may then forget to include them in the medical records. Another explanation for underreporting by medical record review may be that, in the weekly adverse outcomes meeting, as a result of discussions between

surgeons with detailed knowledge of a case, more adverse outcomes may appear including some that were not initially reported. Thus, such a weekly meeting may serve more than one purpose—that is, to learn from the occurrence of adverse outcomes with the purpose of preventing them in the future and to increase the completeness and the accuracy of the adverse outcome reporting information. A solution may be to include a copy of the surgeons' reporting form in the written medical records after the weekly discussion.

Although medical record review may identify more (non-severe) adverse outcomes, its retrospective and outside nature prohibits an easy translation of adverse outcome information in quality improvement initiatives. Routine reporting, in combination with frequent plenary analysis with the aim to learn from experience, provides adequate accuracy for reporting serious adverse outcomes in combination with an ideal setting for quality improvement initiatives. For this reason we believe it to be the best overall method for reporting and prevention of adverse outcomes. Furthermore, this approach of routine reporting in combination with frequent plenary discussion is shown to be generalisable to other hospitals in the Netherlands, including both teaching and non-teaching hospitals.⁷ Such an approach has distinct advantages in comparison with other quality improvement initiatives in surgery such as the National Confidential Enquiry into Patient Outcome and Death (NCEPOD). These advantages include the fact that surgeons can study adverse outcomes in their own population, so they feel responsible for implementing recommendations that result from analysing their data. Data are collected continuously thereby enabling surgeons to study trends over time and to assess whether they are doing better now than in the past. Since adverse outcomes are classified by nature, localisation, and possible determinants and can be linked to surgical procedures, the same doctors who were collecting the data can use these data to support decision making on possible treatment options for patients.

ACKNOWLEDGEMENTS

The authors thank Dr C Jacobi and Dr G H de Bock for their comments on an earlier draft of this paper.

Authors' affiliations

P J Marang-van de Mheen, N van Hanegem, J Kievit, Department of Medical Decision Making, Leiden University Medical Centre, The Netherlands

P J Marang-van de Mheen, J Kievit, Association of Surgeons of the Netherlands, Utrecht, The Netherlands

J Kievit, Department of Surgery, Leiden University Medical Centre, The Netherlands

This study was funded by the Health Research and Development Council (grant number 16450001) and by the Association of Surgeons of the Netherlands.

Conflicts of interest: none.

REFERENCES

- 1 **Brennan TA**, Leape LL, Laird NM, *et al.* Incidence of adverse outcomes and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;**324**:370–6.
- 2 **Gawande AA**, Thomas EJ, Zinner MJ, *et al.* The incidence and nature of surgical adverse outcomes in Colorado en Utah in 1992. *Surgery* 1999;**126**:66–75.
- 3 **Kable AK**, Gibberd RW, Spigelman AD. Adverse outcomes in surgical patients in Australia. *Int J Qual Health Care* 2002;**14**:269–76.
- 4 **Davis P**, Ley-Yee R, Briant R, *et al.* Adverse outcomes in New Zealand public hospitals: principal findings from a national survey, Occasional Paper No 3. New Zealand: Ministry of Health, 2001.
- 5 **Davis P**, Lay-Yee R, Scott A, *et al.* Acknowledgement of “no fault” medical injury: review of patients’ hospital records in New Zealand. *BMJ* 2003;**326**:79–80.
- 6 **Hayward RA**, Hofer TP. Estimating hospital deaths due to medical errors. Preventability is in the eye of the reviewer. *JAMA* 2001;**286**:415–20.
- 7 **Marang-van de Mheen PJ**, Kievit J. Automated registration of adverse events in surgical patients in the Netherlands: the current status (in Dutch). *Ned Tijdschr Geneesk* 2003;**147**:1273–7.
- 8 **Kievit J**, Jeekel J, Sanders FBM. Adverse outcome registration and quality improvement (in Dutch). *Med Contact* 1999;**54**:1363–5.
- 9 **American Society of Anesthesiologists**. New classification of physical status. *Anesthesiology* 1963;**24**:1111.
- 10 **Marang-van de Mheen PJ**, Mertens BJA, van Houwelingen HC, *et al.* Surgery groups differ in adverse outcome probabilities and can be used to adjust hospital comparisons. *J Clin Epidemiol* 2005;**58**:56–62.
- 11 **O’Neill AC**, Petersen LA, Cook F, *et al.* Physician reporting compared with medical-record review to identify adverse medical events. *Ann Intern Med* 1993;**119**:370–6.
- 12 **Wolff AM**, Bourke J, Campbell IA, *et al.* Detecting and reducing hospital adverse events: outcomes of the Wimmera clinical risk management program. *Med J Aust* 2001;**174**:621–5.
- 13 **Firth-Cozens J**. Barriers to incident reporting. *Qual Saf Health Care* 2002;**11**:7.
- 14 **Michel P**, Quenon JL, Sarasqueta AM, *et al.* Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *BMJ* 2004;**328**:199–203.