International Journal of Risk & Safety in Medicine 18 (2006) 65–72
IOS Press

Patient Safety

The epidemiology of medical errors: A few issues in methodology [1]

R.W.M. Giard *
Medisch Centrum Rijnmond-Zuid, Rotterdam, The Netherlands

Abstract. Epidemiological research is needed to determine how many medical errors are made in hospitals. An estimation of error risk requires knowledge of the number of errors made and the total number of patients treated. But how do we define the term 'error'? The outcome of a medical intervention alone is insufficient; the entire treatment process has to be taken into account. For the estimation of error, two different approaches are available: retrospective chart review or prospective ethnographic observation. Both methods should account for processes and outcomes. Both methods have their limitations and can produce measurement errors. In particular, the low reproducibility of the determination of error is worrisome. To achieve safer health care, emphasis should be placed on control of processes, rather than outcomes.

1. Introduction

How many errors are made by doctors? That question needs to be researched systematically. A pioneer was the Scottish surgeon Simpson, who in 1869, analysed which causes, including human factors, determined the surgical mortality during 2098 amputations [18]. Approximately 100 years later Schimmel studied in more than 1000 hospitalisations the incidence of medical error [21].

Much attention was given to the large-scale study of the in-hospital error rate in the American "Harvard medical practice study" of 1991 [6], which was followed by similar surveys in Australia [24], Great Britain [22], Canada [3] and France [15]. Dutch research is in preparation.

In the mentioned analyses it became clear that medical errors are occurring frequently and that they form a structural problem within health care systems.

In 4–17% of hospitalisations errors are made, with personal injury, mortality and in consequence high costs (Table 1). From a professional, political, ethical and economical point of view medical errors are considered as undesirable and research has been demanded in order to improve the functioning of health care along the way of learning and control. The meaning of "learning" in this context is spontaneous action on the basis of experiences and insights.

Increasingly doctors are asked to make the quality of care transparent and therefore useful and reliable data have to be presented by them to become accountable [26].

[1] Originally published as: R.W.M. Giard, De epidemiologie van medische fouten: enkele methodologische kwesties. *Ned Tijdschr Geneeskd* **149** (39) (2005 Sept. 24) 2157-2162. Published with permission of *Ned Tijdschr Geneeskd.*

* Address for correspondence: Dr. R.W.M. Giard, Medisch Centrum Rijnmond-Zuid, locatie Clara, afd. Klinische Pathologie, Postbus 9119, 3007 AC Rotterdam, The Netherlands E-mail: giardr@mcrz.nl.

Table 1
Medical errors as published for various countries

First author (year)	Country	Number of patients		
		Total	Percentage	
			with errors	with preventable errors (in % of number of errors)
Brennan (1991) [6]	US	30,195	3.7	–
Wilson (1995) [24]	Australia	14,179	16.0	51
Vincent (2001) [22]	UK	1014	11.7	48
Baker (2004) [3]	Canada	3745	7.5	37
Michel (2005) [15]	France	8754	5	46

Research into errors confronts us with a number of important methodological questions. How do you define a medical error? How do you develop instruments to examine these problems? How reliable are the outcomes of such research? How should the results be interpreted? What can we do about the problem? Collecting data concerning medical mishaps is never an end in itself, but a means. The aim is the examination and comparison of the quality of medical care, as an incentive for a continuing improvement of it and to assess the impact of the measures taken. In this article I will discuss a number of the mentioned methodological problems more closely.

2. Risk estimate: Numerator and denominator

In quantifying medical error the absolute numbers are of little use and one rather utilizes a relative measure: risk. How large is the chance of an unwanted event per exposure to a medical intervention? The estimate of it is the fraction of the number of accidental events as numerator and as denominator the number of treated patients. That estimate does not have the character of a fixed constant of nature. The numerator includes dissimilar categories of outcomes: mortality, reversible and irreversible personal injury, mental damage and dissatisfaction with respect to the service.

The denominator also consists of several components. The error rate during hospitalisation includes divergent categories of patients, categories which can vary from one institution of care to the other, and the same is true for the spectrum of illness for various disorders or for different wards. The exact meaning of the estimated risk in a certain situation or of the comparison of risk between 2 hospitals depends on the properties of the numerator (which type of risk has been examined; how has it been defined; how was it measured?) and of the denominator (which type of patients was included; which type of treatment and which sickness spectrum are described; how was it measured?).

3. Problem definition

The starting point of a study into the frequency of medical errors is the unambiguously definition of it and that appears to be everything but simple [23]. Several elements have to be distinguished, such as: (a) defining the term “medical error”, (b) arranging medical errors according to different types, (c) classifying events according to their nature, to what extent they are avoidable and if a lack of knowledge or a lack of skill was involved, and (d) differentiating them according to seriousness. We will focus here principally on the first point, the definition of the term “medical error”.

3.1. Conceptual definition

An example of a general definition of the term "error" is the description given by the English psychologist Reason [19]: " 'Error' is used as a generic description of all those situations in which a planned sequence of mental or physical activities do not lead to the intended outcome and where these failures cannot be attributed to a form of chance." Here we have an unwanted outcome as a result of a erroneous process.

Consequently there are two material elements involved with this definition, namely the task that has to be carried out and the result of it, and then there is a normative judgement over this. In medical practice an unwanted outcome – a complication or even death – does not always mean that the doctor could have avoided these outcomes. The trias structure, process and outcome of Donabedian can be used as a conceptual framework for assessing medical actions [4,7]. The relations between structure, process and outcome implies that a medical error is not exclusively defined on the basis of the outcome (Fig. 1) [12]. It is clear that in this definition also an important role is played by the process.

3.2. Complication and error

A complication is every unwanted outcome of medical care, irrespective of the quality of the process, an error means a shortcoming of the quality of the process, irrespective of the outcome [4,7,12]. The relation between these two can be presented in a Venn-diagram (Fig. 2). There can be errors in the process without complications for the patient and complications for the patient without errors in the process. However there are also complications which are the consequence of errors. With all of this a clear answer to the question how a medical error has to be defined has not yet been given: should an error be described primarily starting from the process or from the outcome? If one only looks at the outcome,

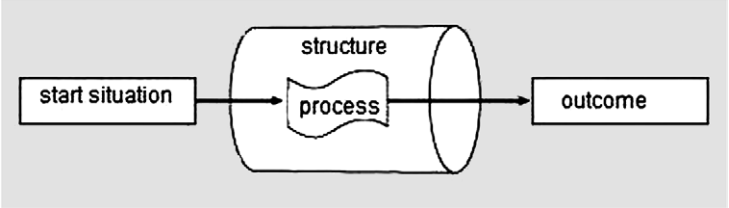

Fig. 1. The so-called trias of Donabedian: structure, process and outcome (12).

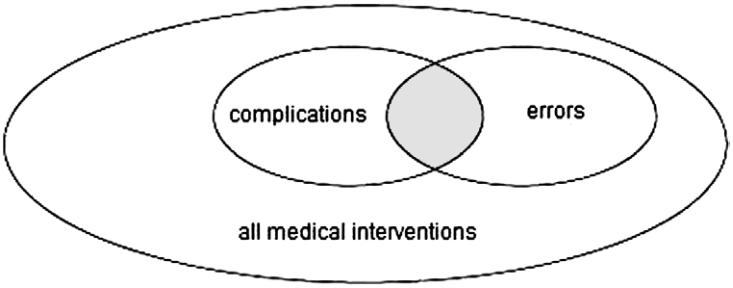

Fig. 2. Venn-diagram in which, within the universe of all medical interventions, populations are situated with unwanted outcomes (complications) and also patients who experienced interventions that were carried out with errors. These populations overlap in part: there can be errors in the process without complications; complications may occur without errors in the process; but there are also complications which are the consequence of errors.

then those errors are missed which do not result in a complication. There are therefore several methods, with which always other patient populations will be identified where improper and wrong medical care took place.

3.3. Operationalisation of the conceptual definition

For the detection of medical errors there are with regard to direction in the time two possibilities: retrospective or prospective analysis. Next comes the question which primary approach is chosen: process – or outcome evaluation [13]. The causal relation between process and outcome is uncertain. Whichever approach one chooses, a practical distinction has to be made between proper and improper medical care, but formulating an univocal and verifiable standard is less simple than it seems.

Take for example the standard with respect to outcomes. When is an inguinal hernia operation been carried out properly? What is an acceptable percentage of recurrences? How many cancers of the breast missed by a screening radiologist can be tolerated? What is the postoperative mortality after coronary bypass operations that is still up to standard? How does one define correct process management and which empirical research is necessary for that? We are going to deal with a complex set of biological, organisational and psychological factors. If one tries to make unambiguous definitions of correct medical actions, one will be faced with the lack of standards based on evidence-based medicine.

4. Retrospective case record review

The majority of studies into medical errors use the retrospective method by means of case record review. Such procedures are time-consuming and therefore costly. First a pallet of criteria is formulated which are used for screening for events that could be indicative for medical errors.

Examples are death, repeated surgery after an earlier intervention, hospital admission after treatment, transfers to intensive care departments, blood transfusions, nosocomial infections, other complaints or injuries which occur during admission, adverse drug events, objections made by the patient or his family or litigious action. Lists of such criteria have been published [3,22,24]. This method is a rather blunt instrument [17]. A practical issue is the fact that one is entirely dependent on the quality of the recordings of considerations and data. Paradoxically, the larger the eagerness to record all information in the file, the larger the chance will be that mistakes are detected. To avoid resulting discrepancies the investigations need to be concentrated on primary data that were defined in advance [13].

If outcome is the primary interest then process errors which do not result in an indicative event will be missed. Such an event can be absent because it did not take place, it was not recognised or it was not written down in the medical record.

Another reason why errors can be missed is that it is possible that they only become apparent after dismissal from the hospital. It has been estimated that as a result of the increasingly shorter duration of admissions and because an increasing number of interventions is done in outpatient departments, a quarter of the complications – the consequence of a medical error or not – are not picked up with retrospective case record review [9].

Because case record review is so time-consuming and nearly all hospitals are nowadays equipped with computerized information systems, alternatives are looked for with which medical errors can be detected in an indirect way [27]. For example excess length of stay, charges, use of medications and mortality can be used as indicators for this purpose.

On the basis of the above it is plausible that the retrospective approach is not a very sensitive method for the detection of errors. It is also a question which value one can attach to the overall percentage of errors for a particular hospital, because within the institution the risks are not evenly spread over the different departments.

5. Prospective observational study of medical errors

Where retrospective case record review generally starts with the appraisal of outcome, prospective research starts with trained observers who first look at the process and then determine what the outcome is and how this outcome relates to the desired result. This approach has been called "participating observation" or "the ethnographic method". In this way an American team examined the work in a department for general surgery and observed serious errors in 18% of the admitted patients [2]. Audit of intensive care departments has taken place in this way as well [8] and at present also such investigations are conducted with TV-cameras which register the activities in for example an operating room, the Intensive Care or the Emergency department [20]. The observations of the reviewers can then afterwards be compared with the television images.

The prospective approach has several advantages in comparison to the retrospective case record review. One is no longer dependent on what has been registered and one is also not dependent on outcomes as possible indicators for process errors. The prospective approach is therefore much more sensitive for the detection of errors. There is the possibility of direct feed back of the observations to the healthcare providers which facilitates interpretation and makes this approach possibly also more specific.

6. Inaccuracies of medical error measurements

6.1. Sensitivity and specificity

How precise and how reproducible are the methods for the study of medical errors? [13] Estimations of precision ask for an independent golden standard, but such a standard does not exist. Thus one must try to make for the different methods of research an estimation as reliable as possible, taking into account variables like the spectrum of sickness and bias. Sensitivity – which percentage of the errors is retrievable? – lies in the order of 70–80% [13,17]. Comparisons of the retro- and prospective methods showed that the prospective approach detected more avoidable errors [16].

The specificity – is an error really an error? – is very important for studies of possible mistakes, because this label may have serious, even legal consequences. Less is known about specificity than about sensitivity but the specificity lies in the order of 95% [16].

6.2. Reproducibility

Of importance is that the agreement between judgements of different observers concerning the question if an event has to be classified as due to an error or not, i.e. reproducibility, mostly appears to be very low [13]. There is also no unanimity with regard to the uncertainty if an unwanted outcome could have been prevented by acting differently [10]. It is exactly this poor reproducibility that forms the Achilles' heel of studies of medical errors: the arbitrariness of the judgement is too large, especially concerning the preventability of errors. To what extent better training or the use of more observers can result in improvement is still insufficiently clear.

6.3. Bias

A another problem concerns bias. If one chooses an unwanted event as starting point for further research of the process, then one has to deal with "hindsight"-bias: ones judgement is generally different – more harsh – as the unfortunate result of the process is known [11]. A second form of bias is observer-bias. The researchers of the study have knowledge of the actors they are going to observe and that may influence their judgement. Anonimisation can be a solution for this problem, however in practice this appeared to be difficult. So it must be concluded that the methodologies of the research and assessment of medical performance have many flaws and that has consequences for the interpretation of the results of this kind of research.

7. The interpretation of research of medical errors

The interpretation of research results can relate to the outcome of one study or it may have to do with the assessment of the meaning of differences between outcomes of several studies. If we examine 100,000 case records for the presence of medical errors with a test method which has a sensitivity of 80% and a specificity of 95%, while in actual fact 1 medical error on the 25 medical interventions occurs (a 4% prevalence of error), then the results can be presented in a tree diagram (Fig. 3). The predictive value of the outcome "medical error" appears then to be only 40% (i.e. $3200/(3200+4800)$). Given the relatively low prevalence of mistakes most of the interventions will not result in any error, but in 5% of them a medical error will be falsely detected. Therefore the consequence will be that the problem of medical errors is over-estimated, because in this case the measured error rate is 8% (i.e. $(3200+4800)/100{,}000$).

In the table the comparison of outcomes of several studies is shown (Table 1). Are we dealing here with differences in quality of care? Several explanations for such differences do exist [14]. We might be looking at differences in definitions and differences in methods of assessment. Furthermore outcomes of interventions are not only determined by the quality of care, but also by numerous other patient – and environment related factors. Moreover the composition of the patient population ("casemix") may vary from hospital to hospital. Also chance may play a role.

Before a certain error rate can be seen as proof of poor quality or that it can be concluded that in one hospital more errors are made than in an other, all possible factors that might be of influence for those outcomes and all alternative explanations have to be taken into account [5].

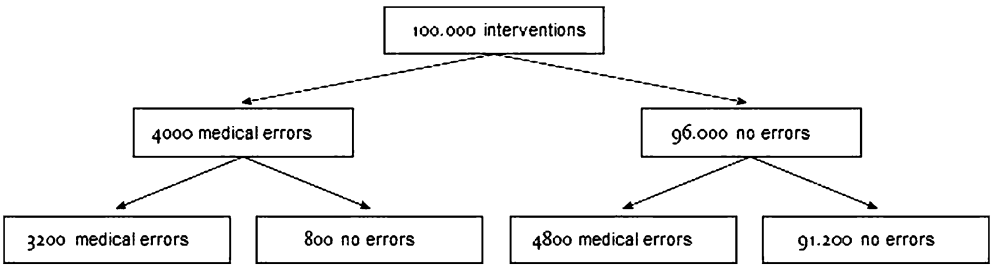

Fig. 3. Tree diagram according to Bayes for 100,000 medical interventions. The middle row shows the real situation; the bottom row shows the results of a study of medical errors with a sensitivity of 80% and a specificity of 95%.

8. Towards a safer health care by means of learning and control

It is tempting to consider all medical errors as unacceptable, but that wouldn't be realistic. Shouldn't we question the possibility of having an absolutely safe health care system? [1] The purpose of epidemiological studies of errors is not only descriptive, but also normative: to assess if the given care is conform the formulated requirements. However, are those requirements so univocal and verifiable? What is the realisable upper limit of medical performance that can be realised?

The burning question is: what is the practical meaning of the epidemiological data about medical errors and how can we use those data to realise a safer health care? Here the question is at hand, should we give priority to outcome – or to process assessment. The unreliable relation between processes and outcomes limits the usefulness of outcome evaluations as an instrument for quality assessment. The largest advance can be realised by means of process evaluation; however this approach does not only need a medical–technical analysis but also a broad organisational context [1,13]. Nevertheless, to evaluate the effects of process improvements we cannot do without systematically collected data concerning the outcome of medical interventions, knowing that the methodologies that are available are still suffering from growing pains [13,17].

9. Conclusion

The media show huge eagerness regarding percentages of medical errors which are presented as "shocking", however such publications create unfounded mistrust towards doctors, resulting in irritation and frustration of health professionals [26]. Accountability means that performances are compared with a standard; therefore the first step in this process of answerability is the conception of normative guidelines along the way of evidence-based medicine. Such an effort can be expected, but it is made a difficult task especially by the complexity of contemporary medicine [25]. Maybe less errors will be made if the handling of safety is approached differently, namely in the way of process management. However, that choice has to be made on the basis of rational arguments and demands various empirical studies and the exploration of the current barriers on the way towards maximum safety [1].

Competing interests: None declared. Financial support: None declared.

References

[1] R. Amalberti, Y. Auroy, D. Berwick and P. Barach, Five system barriers to achieving ultrasafe health care, *Ann. Intern. Med.* **124** (2005), 756–764.
[2] L.B. Andrews, C. Stocking, T. Krizek, L. Gottlieb, C. Krizek, T. Vargish et al., An alternative strategy for studying adverse events in medical care, *Lancet* **349** (1997), 309–313.
[3] G.R. Baker, P.G. Norton, V. Flintoft, R. Blais, A. Brown, J. Cox et al., The Canadian adverse events study: the incidence of adverse events among hospital patients in Canada, *CMAJ* **170** (2004), 1678–1686.
[4] J.B. Battles and R.J. Lilford, Organizing patient safety research to identify risks and hazards, *Qual. Saf. Health Care* **12**(Suppl. 2) (2003), ii2–7.
[5] S.M. Bird, D. Cox, V.T. Farewell, H. Goldstein, T. Holt and P.C. Smith, Performance indicators: good, bad, and ugly, *J. R. Stat. Soc. Ser. A* **168** (2005), 1–27.
[6] T.A. Brennan, L.L. Leape, N.M. Laird, L. Hebert, A.R. Localio, A.G. Lawthers et al., Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard medical practice study I, *N. Engl. J. Med.* **324** (1991), 370–376.
[7] A. Donabedian, Evaluating the quality of medical care, *Milbank Mem. Fund Q* **44** (1966), 166–206.

[8] Y. Donchin, D. Gopher, M. Olin, Y. Badihi, M. Biesky, C.L. Sprung et al., A look into the nature and causes of human errors in the intensive care unit, *Crit. Care Med.* **23** (1995), 294–300.
[9] A.J. Forster, H.D. Clark, A. Menard, N. Dupuis, R. Chernish, N. Chandok et al., Adverse events among medical patients after discharge from hospital, *CMAJ* **170** (2004), 345–349.
[10] R.A. Hayward, T.P. Hofer, Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer, *JAMA* **286** (2001), 415–420.
[11] K. Henriksen and H. Kaplan, Hindsight bias, outcome knowledge and adaptive learning, *Qual. Saf. Health Care* **12**(Suppl. 2) (2003), ii46–50.
[12] T.P. Hofer, E.A. Kerr and R.A. Hayward, What is an error?, *Eff. Clin. Pract.* **3** (2000), 261–269.
[13] R.J. Lilford, M.A. Mohammed, D. Braunholtz and T.P. Hofer, The measurement of active errors: methodological issues, *Qual. Saf. Health Care* **12**(Suppl. 2) (2003), ii8–12.
[14] J. Mant. Process versus outcome indicators in the assessment of quality of health care, *Int. J. Qual. Health Care* **13** (2001), 475–480.
[15] P. Michel, J.L. Quenon, A. Djihoud, S. Tricaud-Vialle, A.M. de Sarasqueta and S. Domecq, Les événements indésirables graves liés aux soins observés dans les établissements de santé: premiers résultats d'une étude nationale, *Études et résultats* **398** (2005), 1–15.
[16] P. Michel, J.L. Quenon, A.M. de Sarasqueta and O. Scemama, Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals, *BMJ* **328** (2004), 199.
[17] G. Neale and M. Woloshynowych, Retrospective case record review: a blunt instrument that needs sharpening, *Qual. Saf. Health Care* **12** (2003), 2–3.
[18] D. Neuhauser, Surgical experience, hospital size and severity adjusted mortality: James Y. Simpson, 1869, *Qual. Saf. Health Care* **14** (2005), 67–68.
[19] J. Reason, *Human Error*, Cambridge University Press, Cambridge, 1990, p. 9.
[20] D.A. Rogers, G. Regehr and J. MacDonald, A role for error training in surgical technical skill instruction and evaluation, *Am. J. Surg.* **183** (2002), 242–245.
[21] E.M. Schimmel, The hazards of hospitalisation, *Ann. Intern. Med.* **60** (1964), 100–110.
[22] C. Vincent, G. Neale and M. Woloshynowych, Adverse events in British hospitals: preliminary retrospective record review, *BMJ* **322** (2001), 517–519.
[23] S.N. Weingart, L.I. Iezzoni, Looking for medical injuries where the light is bright, *JAMA* **290** (2003), 1917–1919.
[24] R.M. Wilson, W.B. Runciman, R.W. Gibberd, B.T. Harrison, L. Newby and J.D. Hamilton, The quality in Australian health care study, *Med. J. Aust.* **163** (1995), 458–471.
[25] T. Wilson, T. Holt and T. Greenhalgh, Complexity science: complexity and clinical care, *BMJ* **323** (2001), 685–688.
[26] Wetenschappelijke Raad voor het Regeringsbeleid. Bewijzen van goede dienstverlening. Hfdst 7, Amsterdam University Press, Amsterdam, 2004, pp. 185–216.
[27] C. Zhan and M.R. Miller, Excess length of stay, charges, and mortality attributable to medical injuries during hospitalisation, *JAMA* **290** (2003), 1868–1874.