

What makes an error unacceptable? A factorial survey on the disclosure of medical errors

DAVID L. B. SCHWAPPACH AND CHRISTIAN M. KOECK

University Witten/Herdecke, Health Policy & Management, Witten, Germany

Abstract

Background. Although the importance of disclosing medical errors to patients has been argued, little is known about the relative effect of different attributes of error handling and communication on patients' judgments about errors.

Objectives. This study investigates how different characteristics of medical errors and of physicians' subsequent handling of errors contribute to patients' evaluations of the incident and their attitudes towards potential consequences and sanctions for the physician.

Materials and methods. A factorial survey using the vignette technique presented hypothetical scenarios involving medical errors to members of the general public in an Internet-based study. Members of a German Internet survey panel participated ($n = 1017$). Multiple ordered logistic regression models were estimated to explain citizens' judgments of error severity and their attitudes towards reporting of errors, wishing for referral to another physician, and supporting sanctions against the health professional involved as a response to characteristics of the presented errors.

Results. While the severity of the outcomes of errors remains the most important single factor in the choice of actions to be taken, the professional's approach to the error is regarded as essential in the overall evaluation of errors and the consideration of consequences. In errors with a severe outcome, an honest, empathic, and accountable approach to the error decreases the probability of participants' support for strong sanctions against the physician involved by 59%. Judgments were only marginally affected by respondents' characteristics.

Conclusions. The handling of errors strongly contributes to citizens' choice of actions to be taken, and they are sensitive to failures to name the incident as an 'error'. For the success of de-individualized, systems-oriented approaches to errors, communication of clear accountability to patients will be crucial.

Keywords: factorial survey, Internet, malpractice, medical errors, survey analysis

Medical errors are among the most serious quality problems in health care systems and are associated with considerable health-related harm and economic burden [1]. Medical errors also have the potential to compromise the physician–patient relationship and to undermine patients' trust in the health care system as a whole. 'To err is human', and even if all effective measures were taken to prevent errors, the occurrence of some errors has to be accepted as an inevitable side-effect of medical care. Besides the prevention of errors, it is therefore essential that health care professionals approach errors in a way that allows patients to preserve confidence in the health care system. A number of authors have stressed the importance of disclosure of errors to patients, the role of apologizing, and the need for honest and caring communication [2,3]. Empirical studies have shown that patients have strong preferences

towards the disclosure of errors and the provision of information about underlying causes, consequences, and prevention of recurrences [4,5]. There is also evidence that the way the incident is handled by health care professionals influences whether and which consequences are drawn by affected patients (or their families) [6]. If the trauma of being harmed by a health care professional is followed by the trauma of not being taken seriously and communicated with sensitively, the likelihood of taking legal actions against the involved persons increases [7,8]. However, very little is known about how different characteristics of a medical error, such as outcome severity, and of the subsequent approach to this error contribute to patients' evaluations of the incident and their attitudes towards potential consequences. We conducted a factorial survey using the vignette technique [9] that presented

Address reprint requests to David L. B. Schwappach, Department of Health Policy and Management, Faculty of Medicine, Alfred-Herrhausen-Str. 50, 58448 Witten, Germany. E-mail: davids@uni-wh.de

hypothetical scenarios involving a medical error to members of the general public. The main objective was to investigate how characteristics of the error itself and the way the error is handled relate to citizens': (i) judgments of error severity; (ii) wishes to be referred to another physician; (iii) attitudes towards reporting of errors; and (iv) requests for sanctions for the involved health professional.

Methods

Survey instrument

The survey was administered through the Internet as part of a larger survey study about citizens' attitudes towards errors in Germany. The part of survey we report on here presented vignettes to participants that described a hypothetical scenario of a patient experiencing a medical error, and excerpts from the physician's communication about the incident. Each vignette was followed by the same four questions. These were:

1. To rate the severity of the incident on a seven-point likert scale ranging from 'minor error' to 'very severe error';
2. Whether participants would seek referral to another physician if they had experienced the described situation, with response codes 'no', 'rather no', 'rather yes', 'yes', and 'I don't know';
3. Whether participants would report the error (e.g. to the sickness fund or board of physicians), with response codes 'yes', 'no', and 'I don't know';
4. Whether the error or the handling of the error should have consequences for the physician, and which sanctions would be appropriate, with response codes 'no consequences', 'moderate consequences' (e.g. admonition by the board of physicians, obligation to further training), and 'major consequences' (e.g. dismissal, revocation of the license to practice, lawsuit). Throughout the survey, 'errors in the health care system' were defined as: '... an avoidable injury caused by medical management rather than the underlying condition of the patient'. Examples of errors and negative outcomes not involving errors were given. These explanations were available on all survey pages through hyperlinks.

Vignettes

Each vignette contained a combination of six attributes that were derived from the literature. These were: the severity of the initial condition; the type of error; the severity of consequences associated with the error; elements of physician-sided communication about the error; whether the error is acknowledged as such; and whether the physician apologizes. Vignettes differ in the levels of these attributes and differences in levels are used to explain variance in respondents' judgments. Attributes, assigned levels, and their translation into vignette phrases are summarized in the Appendix (A1). The number of attributes and levels results in 486 possible combinations and versions of the vignette. The experimental

design software SPEED was used to reduce this number in a fractional factorial design [10]. The main-effects design has an orthogonal structure, i.e. ensures the absence of multicollinearity and minimizes mis-specification [11]. The resulting 27 combinations were then translated into vignette stories and paired in nine sets, each containing three vignettes. An example vignette is provided in the Appendix (A2). The survey was programmed so that each participant was randomly assigned to one of the nine sets, and the ordering of vignettes among sets was randomly rotated to avoid ordering effects.

Survey administration

The survey was administered through the 'GesundheitsPanel' (available online at <http://www.gesundheitspanel.de>), an Internet survey panel that was initiated in 2003. Details of the panel have been described elsewhere [12]. In brief, the GesundheitsPanel provides the environment under which specific survey projects can be processed, fielded, and administered to samples of panelists. Participants are mainly recruited via non-electronic media advertisements and reports, and presentations at adult education centers. To become a panel member and to be eligible for participation in future surveys, interested citizens are first invited to participate in a master survey to collect detailed demographic and health-related data. Panelists agree to be surveyed regularly, and consent to membership and data protection rules. Samples of panel members are drawn, invited by e-mail to participate, and specific survey projects are then administered to these subjects online via personalized, secure Internet pages. Finally, survey responses are linked to the master data available for each respondent. Since the panel is explicitly advertised as a health care panel, membership is most attractive to patients, their relatives, persons interested in the German health care reform, and health care professionals (who are not actively recruited). The panel therefore reflects the views and attitudes of those with an intense background of personal experiences of the health care system.

Data analysis

Responses to vignettes were treated as separate observations. The unit of analysis was the judgment provided to each question in response to a vignette, and not the individual respondent. Four separate logistic and ordered logistic regression models were estimated, in which vignette attributes and participants' characteristics (independent variables) explained the binary (question 3) or ordinal responses (questions 1, 2, and 4). Vignette characteristics were included as sets of dummy variables. In addition, the following factors were considered as explanatory variables: respondents' age, sex, education, income, occupation, and place of residence, whether they work as a health care professional, perceived health, type of health insurance, and utilization of the Internet. Responses to two survey questions were also included: whether respondents had experienced a medical error in their own care (yes/no) and whether they are afraid of experiencing errors (no/sometimes/frequently). Results of the regression models are presented as odds

ratios (OR). In ordered logistic regression, the OR is interpreted as the odds of observing a response in a higher outcome category versus the lower categories for a unit change in the independent variable. The value of the OR does not depend on the value of the category ('proportional odds assumption').

For each model, responses in the opt-out category were dropped from analysis. The sample for each regression model is therefore defined as the number of respondents multiplied by the number of vignettes, minus the number of participants who opted out, minus those with missing values for any of the independent variables. Because some questions in the master survey allow response refusal, there were missing data in personal characteristics of some participants. These were imputed using single Hotdeck imputation [13]. Responders and non-responders were compared using paired *t*-tests and chi-square tests for nominal variables. For all statistical tests, a *P*-value of 0.05 was considered significant. Confidence intervals are reported on the 95% level. Data were analyzed with STATA v8 [14].

Model specification

Each participant responded to three vignettes, therefore responses provided by each individual cannot be regarded as independent. All regression models were therefore specified with Huber–White sandwich variance estimators for clustered data, with responders as unit of clustering. The cluster variance estimator is robust to mis-specification and within-cluster correlation [15]. From the beginning of data analysis, it was obvious that there were only small numbers of observations in the bottom levels 1–4 (starting with 'minor error') of the original seven-point error severity rating scale in question 1. A generalized Hausmann-type test was then performed, and confirmed that the low frequency categories (levels 1–4) could be condensed into one category in the empirical analysis. Based on these diagnostics and for easier interpretation of results, the final ordered logistic model explains variance in judgments on a four-point scale. All models were tested for multicollinearity, mis-specification, and omitted variables. The ordered logistic models were also tested for violations of the parallel regression assumption using a Brant test [16].

Results

Of the 1200 persons invited, 1017 participated (response rate 85%) and 984 (82%) completed the questionnaire. Details of respondents' characteristics are presented in Table 1. Compared with responders, non-responders were slightly younger (43 compared with 45 years; $P < 0.001$), more likely to be health care professionals (25% compared with 12%; $P < 0.001$), and less likely to be female (41% compared with 59%, $P < 0.001$). Thirty percent of respondents reported that they had experienced an error in their own care. Nearly 70% of these participants provided extensive descriptions of their experiences, indicating that the majority did indeed have specific events in mind when answering. Many respondents stated that they feared experiencing an error in their own care sometimes, e.g. in a specific treatment situation (58%), or even frequently (16%).

Table 1 Respondents' characteristics

Personal characteristic	Number of responders (%)
Mean age (years)	45
<25	36 (4)
25–34	199 (20)
35–44	300 (30)
45–59	348 (34)
≥60	134 (13)
Gender (% female)	597 (59)
Health care professionals	124 (12)
Years of education, including primary school (years)	
≤12	199 (20)
13–15	229 (23)
16–18	300 (30)
≥19	289 (28)
Disposable income (€ /month) ¹	
<1100	358 (35)
1100–2600	492 (48)
>2600	167 (16)
Occupation	
Full time	493 (48)
Part time	158 (16)
Unemployed	38 (4)
Homemaker	61 (6)
Student	97 (10)
Retired	170 (17)
Perceived health	
Very good	163 (16)
Good	438 (43)
Fair	277 (27)
Bad	120 (12)
Very bad	19 (2)
Statutory health insurance (%)	860 (85)
Initial begin of Internet utilization	
<12 months	78 (8)
1–3 years	318 (31)
>3 years	621 (61)
Average weekly Internet use	
<2 hours per week	175 (17)
2–10 hours per week	542 (53)
>10 hours per week	300 (30)

Total $n = 1017$.

¹Number of imputed values: occupation = 37; income = 70.

The frequency distributions of responses calculated over all vignettes (Table 2) show that the design of vignettes stimulated responses sensitive to vignette characteristics and provoked variability in participants' judgments. There was no evidence of mis-specification or multicollinearity in the regression models. Results of the final regression models are provided in Tables 3–6. Four of the six vignette attributes were significant in all regression models. Type of error, severity of error, acknowledgment of error, and apologizing for the error

Table 2 Frequency distribution of responses to questions, over all vignettes

Questions and response codes ¹	Responses (%)
Question 1: error severity rating	<i>n</i> = 2924 ²
Very severe error	25.9
Rather severe error	19.5
Rather minor error	30.3
Minor error ³	24.3
Question 2: seeking referral to another physician	<i>n</i> = 2840
Yes	38.9
Rather yes	24.4
Rather no	25.1
No	11.6
Question 3: reporting the error	<i>n</i> = 2466
Yes	56.9
No	43.1
Question 4: sanctions for the physician	<i>n</i> = 2646
None	28.3
Moderate	57.5
Major	14.1

¹See Methods for original wording of questions.

²*n* differs between questions due to drop out during survey progress and due to opting out in particular survey questions.

³Retrospectively summarized from four levels.

strongly influenced participants' judgments. In all models, the severity of the error was the most important single variable, with ORs ranging from 13 to 40 in the four models for 'severe outcome of error' as compared with 'mild outcome'. Participants were very sensitive to the details the vignette reported about the acknowledgment of error. While the unambiguous acknowledgment that an error has happened was evaluated positively and decreased the probability of negative judgments in all models, equivocal statements suggesting the patient's share of responsibility had no effect, or even increased the probability of negative judgments (Table 4). Compared with the 'narrow' communication style, only the 'comprehensive' style significantly affected respondents' responses, and slightly decreased the probability of negative evaluations in all but the 'reporting model'.

There was no systematic effect of participants' personal characteristics on their judgments. Gender, disposable income, and occupation were not related to respondents' judgments in either model. Higher educated responders rated the described error as less severe and were less likely to opt for reporting of the error. Surprisingly, longer experience in personal Internet utilization was only associated with a higher probability of seeking referral to another physician. Citizens of larger municipalities compared with rural areas were also more likely to prefer referral. Unexpectedly, having experienced an error in one's own care did not affect respondents' appraisal, but being afraid of errors both 'sometimes' or 'frequently' significantly contributed to the probability of negative judgments in all models. Many participants provided extensive

Table 3 Vignette predictors of error severity rating (robust ordered logit regression)

Variable	Odds ratio	95% CI
Vignette set	1.03	0.99–1.07
Severity of initial condition (base category 'mild')		
Moderate	0.92	0.76–1.11
Severe	1.27	1.05–1.53
Type of error (base category 'oversight of diagnostic information')		
Oversight of allergy	2.02	1.72–2.36
Wrong treatment	1.15	0.99–1.33
Severity of error (base category 'mild')		
Moderate	2.99	2.56–3.50
Severe	23.04	18.83–28.20
Acknowledgment of error as such (base category: blank, no information provided)		
Yes	0.78	0.66–0.93
No	0.96	0.81–1.13
Communication (base category 'narrow')		
Basic	1.01	0.87–1.19
Comprehensive	0.80	0.68–0.94
Apology (base category: blank, no information provided)		
Yes	0.77	0.66–0.90
Having experienced a medical error in one's own care		
Yes	1.02	0.84–1.25
Being afraid of experiencing medical errors in one's own care (base category 'no')		
Sometimes	1.35	1.08–1.70
Frequently	1.82	1.31–2.52
Age (years)	1.01	1.00–1.02
Male	1.07	0.87–1.32
Health professional	0.70	0.54–0.90

CI, confidence interval.

Total *n* (judgments) = 2889; total *N* (individuals) = 966. McKelvey and Zavonia's $R^2 = 0.35$.

feedback. Most comments related to the physician's approach to the error and either acknowledged honesty and courage or criticized disrespect and self-righteousness. Several participants argued that their general impression of the quality of care provided would also be an important aspect and many noted that their judgment would also depend on whether the physician had been reported because of errors previously.

Health care professionals differed from other respondents in all models. They provided lower error severity ratings, and were less likely to prefer reporting the error, wishing referral to another physician, or requesting sanctions. In their qualitative

Table 4 Vignette predictors of ‘wish for referral’ (robust ordered logit regression)

Variable	Odds ratio	95% CI
Vignette set	1.04	1.00–1.08
Severity of initial condition (base category ‘mild’)		
Moderate	0.79	0.64–0.96
Severe	1.13	0.93–1.38
Type of error (base category ‘oversight of diagnostic information’)		
Oversight of allergy	2.17	1.85–2.56
Wrong treatment	1.28	1.09–1.50
Severity of error (base category ‘mild’)		
Moderate	2.17	1.85–2.55
Severe	12.92	10.61–15.74
Acknowledgment of error as such (base category: blank, no information provided)		
Yes	0.44	0.36–0.52
No	1.27	1.06–1.52
Communication (base category ‘narrow’)		
Basic	1.01	0.85–1.22
Comprehensive	0.55	0.46–0.65
Apology (base category: blank, no information provided)		
Yes	0.73	0.62–0.86
Having experienced a medical error in one’s own care		
Yes	1.11	0.91–1.36
Being afraid of experiencing medical errors in one’s own care (base category ‘no’)		
Sometimes	1.33	1.07–1.66
Frequently	1.86	1.36–2.54
Age (years)	0.99	0.98–1.00
Male	1.03	0.84–1.27
Health professional	0.62	0.47–0.82

CI, confidence interval.

Total *n* (judgments) = 2807; total *N* (individuals) = 958. McKelvey and Zavonia’s $R^2 = 0.33$.

comments, many participants with a health care background doubted whether the vignette described an error or rather an ‘unfortunate side-effect or incident’. In the vignettes reporting oversight of allergy as type of error, health care professionals often argued it was ‘the responsibility of the patient to check whether prescribed drugs may have allergenic potential’.

We also computed best-/worst-case scenarios for each model, comparing the probability of negative judgments following the best versus the worst possible approach to the error. The only difference between the scenarios was the communication style (comprehensive versus narrow), the acknowledgment of error as such (‘yes’ versus ‘no’ for infor-

Table 5 Vignette predictors of ‘reporting of error’ (robust logit regression)

Variable	Odds ratio	95% CI
Vignette set	1.07	1.01–1.12
Severity of initial condition (base category ‘mild’)		
Moderate	0.90	0.68–1.19
Severe	1.34	1.05–1.82
Type of error (base category ‘oversight of diagnostic information’)		
Oversight of allergy	1.97	1.58–2.47
Wrong treatment	0.99	0.80–1.24
Severity of error (base category ‘mild’)		
Moderate	3.51	2.81–4.38
Severe	39.56	29.06–53.86
Acknowledgment of error as such (base category: blank, no information provided)		
Yes	0.71	0.55–0.92
No	1.01	0.79–1.31
Communication (base category ‘narrow’)		
Basic	1.28	0.97–1.68
Comprehensive	0.83	0.64–1.07
Apology (base category: blank, no information provided)		
Yes	0.73	0.58–0.91
Having experienced a medical error in one’s own care		
Yes	1.00	0.77–1.30
Being afraid of experiencing medical errors in one’s own care (base category ‘no’)		
Sometimes	1.70	1.26–2.31
Frequently	2.55	1.69–3.85
Age (years)	0.99	0.97–1.00
Male	1.12	0.85–1.46
Health professional	0.36	0.24–0.54

CI, confidence interval.

Total *n* (judgments) = 2443; total *N* (individuals) = 927. McKelvey and Zavonia’s $R^2 = 0.45$.

mation provided), and whether the physician apologizes (‘yes’ versus ‘no’ for information provided). Assuming a severe error, and keeping all variables other than these ‘handling’ variables constant, the predicted probability of observing a response in the ‘extreme severe’ error rating category decreased from 0.51 to 0.34 with the optimal approach to the error. The probability of wishing referral to another physician decreased from 0.74 to 0.33, that of reporting the error decreased from 0.92 to 0.83, and that of observing a response in the ‘major consequences for the physician’ category decreased from 0.36 to 0.15, which equals a relative decrease of 59%. An optimal

Table 6 Vignette predictors of ‘support for sanctions’ (robust ordered logit regression)

Variable	Odds ratio	95% CI
Vignette set	1.04	0.99–1.08
Severity of initial condition (base category ‘mild’)		
Moderate	0.93	0.75–1.17
Severe	1.12	0.93–1.44
Type of error (base category ‘oversight of diagnostic information’)		
Oversight of allergy	1.97	1.62–2.39
Wrong treatment	1.32	1.10–1.58
Severity of error (base category ‘mild’)		
Moderate	2.78	2.29–3.38
Severe	31.76	24.29–41.53
Acknowledgment of error as such (base category: blank, no information provided)		
Yes	0.53	0.43–0.64
No	0.91	0.74–1.12
Communication (base category ‘narrow’)		
Basic	1.17	0.97–1.42
Comprehensive	0.79	0.65–0.96
Apology (base category: blank, no information provided)		
Yes	0.75	0.63–0.90
Having experienced a medical error in one’s own care		
Yes	1.00	0.81–1.26
Being afraid of experiencing medical errors in one’s own care (base category ‘no’)		
Sometimes	2.16	1.67–2.80
Frequently	2.69	1.89–3.84
Age (years)	1.00	0.98–1.01
Male	1.13	0.90–1.42
Health professional	0.53	0.37–0.75

CI, confidence interval.

Total n (judgments) = 2623; total N (individuals) = 936. McKelvey and Zavonia’s $R^2 = 0.42$.

approach to the error has even more weight in the ‘moderate error’ scenario. Compared with the ‘worst approach to the error’, the relative decrease in the predicted probabilities of an ‘extreme severe’ error rating in the best case scenario was 49%, 76% in the probability for wishing referral, 39% for reporting the error, and 77% for requesting ‘major consequences’.

Discussion

To the authors’ knowledge, this is the first quantitative study to investigate the relative importance of several aspects of

medical errors and their handling on laypersons’ evaluation of the error. As others, we observed that the professional’s approach to the error is regarded as essential in the overall evaluation of errors and the consideration of consequences [6,7,17]. We also provide further evidence that full disclosure of errors may strengthen rather than undermine the relationship between patients and physicians [18]. In particular, in errors with moderate negative outcomes, the honest approach to the mistake actually decreases the likelihood of the patient reporting the error and the desire to punish the involved professional. Still, the severity of the outcome of errors remains the most important single factor in the choice of actions to be taken. In concordance with the study by Blendon *et al.*, members of the general public are well prepared to discriminate according to health outcomes and support strong sanctions towards individual health professionals if patients are severely harmed, or even die [19]. An important and novel finding of our study is that the unequivocal statement that an error has happened, and taking on responsibility, has at least as much preventive effect on the advocating of sanctions as apologizing. However, professionals commonly avoid the naming of the incident. Gallagher *et al.* reported on a focus group study among US physicians, saying that many professionals ‘choose their words carefully’, which most often involves mentioning the adverse event but not explicitly stating that an error took place [4]. In our study, participants were very sensitive to precisely such attempts to mislead patients and did not accept professionals ‘muddling through the issue elegantly’.

As in the Agency for Health Care Research and Quality survey, a considerable fraction of participants in our study was concerned about experiencing errors in their own care, and these responders were also more likely to provide negative judgments [20]. Future studies may investigate whether being concerned about errors and being more willing to respond to them have a causal relationship, coincide with, or are both outcomes of a more general critical attitude towards the health care system.

The finding that the moderate compared with the mild initial condition has no or even counterintuitive impact on respondents’ judgments is concerning. Although we pilot tested descriptions of attributes in small samples of students and panelists in ranking exercises and with series of vignettes that contained only differences in levels of one attribute, it seems likely that the mild versus the moderate health states suffered discriminatory power. It should be noted though that every respondent received only three out of the 27 vignettes. The number of different levels per characteristic presented to each individual is therefore very small. As a consequence, the potential lack of differentiation occurred mainly between, and not within, subjects’ judgments.

This study also has some limitations: The first regards the sample, which is biased towards the well educated, those with intensive experiences of health care, and those using the Internet. Health professionals are over-represented in the sample. It can be argued that those persons with direct and frequent contact with the health care system, either as patients, relatives, or professionals, have a more realistic idea of the situations in which errors happen and have less

problems with identifying and seriously judging the hypothetical vignettes. On the results level, respondents' judgments were only marginally and unsystematically affected by personal characteristics and Internet utilization. In addition, the fraction of responders reporting errors in their own care is comparable to that of other studies. According to Blendon *et al.*, the fraction of respondents reporting errors in their own care during the past 2 years ranged between 18% in the UK and 28% in the US [21]. About 42% of participants in the National Patient Safety Foundation survey reported they had been involved, either personally or through a friend or relative, in a situation where a medical mistake was made [22]. In the only study presenting German data, a representative survey undertaken in 2002 by a major German statutory health insurance organization, 20% of respondents reported having experienced 'medical malpractice' in their own care [23]. Given that our broader definition of medical errors also included events in which health care professionals other than physicians were involved, we did not limit experiences to a specific time frame, and panelists are likely to be at greater risk for experiencing medical errors, our figure does not seem artificially high. Finally, compared with stand-alone Internet surveys, the problem of self-selectivity is alleviated within the panel design, since members selected themselves for membership, but not for participation in this survey [24]. Still, members of the survey panel probably differ from the general population in that they are more likely to share a proactive attitude. While this may have systematically affected the overall anticipated preparedness to take actions, it seems rather unlikely that the relative effects of the error and handling attributes are also biased.

Three limitations relate to the vignette approach and the presentation of hypothetical cases to participants. Firstly, respondents were provided the de facto details of the error, had a bird's eye view of the professional's approach to the error, and judged this against the 'whole truth'. In reality, however, patients have to judge a physician's communication against their own vague interpretation of an event. In conclusion, for the vignettes describing dishonest communication, respondents judged the attempt to deceive that they had discovered. However, this approach mirrors one of the main arguments in the literature on 'disclosing bad news' to patients. Namely, that it is the revealing of deception that erodes patients' trust and that patients' 'sense of betrayal will probably far outweigh any distress from being told the truth' [25,26]. Secondly, whether the judgments made in a 'virtual setting' would be confirmed by real actions if respondents themselves actually experienced the error remains unclear. Discrepancies between social judgments in vignette surveys and actual behavior have been reported [27]. Social desirability bias towards an active and autonomous patient may have lured participants into an overestimation of their own response to errors. Finally, we provided only written, albeit detailed, information in the vignettes. Other studies have shown that how a physician says something may be as important as what they say, in respect to a patient's litigious attitude [28]. Non-verbal aspects of communication were not presented in our study and we do not know how participants filled this missing information in [29].

Overall, participants were rather reserved towards advocating strong sanctions against physicians and the results of this study demonstrate the significance of error management. Most patients accept that errors are not entirely preventable, but they expect accountability and clear words. These clear words should include the acknowledgment that something wrong has happened, that measures will be taken to prevent future events, the option to be referred to a different provider, and an expression of sincere regret. The relabeling of errors as 'unhappy events' should be strongly avoided since it irritates patients' trust in their perceptions and increases distress.

Recently, the quality improvement approach to medical errors, which requires systems thinking and establishment of a culture of organizational learning rather than individual 'naming and shaming', has been advocated more and more [30]. These developments seem promising for learning through, and finally preventing errors. However, it will be crucial that this approach to errors among the medical profession is not being misused to negate patients' needs in the handling and disclosure of errors. It would be cynical to expect injured patients to understand errors as a 'chance for organizational learning'. Their needs to identify and blame liable individuals are a legitimate response to harm, and failure to assign blame when it is due is undesirable [31]. To become successful, efforts to deindividualize error origin and responsibility and to promote errors as system failures must be accompanied by a strategy for clear naming and explanation of the incident to patients, and accountability of the organization as a whole must transfer onto individuals identifiable by patients.

Acknowledgements

We would like to thank two anonymous referees for their very valuable comments and suggestions, which helped us to improve the paper. The participation of panel members in this study is greatly acknowledged.

References

1. von Laue NC, Schwappach DL, Koeck CM. The epidemiology of medical errors: a review of the literature. *Wien Klin Wochenschr* 2003; **115**: 318–325.
2. Hebert PC. Disclosure of adverse events and errors in healthcare: an ethical perspective. *Drug Saf* 2001; **24**: 1095–1104.
3. Vincent CA, Coulter A. Patient safety: what about the patient? *Qual Saf Health Care* 2002; **11**: 76–80.
4. Gallagher TH, Waterman AD, Ebers AG, Fraser VJ, Levinson W. Patients' and physicians' attitudes regarding the disclosure of medical errors. *JAMA* 2003; **289**: 1001–1007.
5. Hingorani M, Wong T, Vafidis G. Patients' and doctors' attitudes to amount of information given after unintended injury during treatment: cross sectional, questionnaire survey. *Br Med J* 1999; **318**: 640–641.

6. Witman AB, Park DM, Hardin SB. How do patients want physicians to handle mistakes? A survey of internal medicine patients in an academic setting. *Arch Intern Med* 1996; **156**: 2565–2569.
7. Vincent C, Young M, Phillips A. Why do people sue doctors? A study of patients and relatives taking legal action. *Lancet* 1994; **343**: 1609–1613.
8. Wu AW. Handling hospital errors: is disclosure the best defense? *Ann Intern Med* 1999; **131**: 970–972.
9. Rossi PH, Nock SL. *Measuring Social Judgments. The Factorial Survey Approach*. London: SAGE Publications, 1982.
10. Bradley M. *Users' Manual for SPEED Version 2.1 Stated Preference Editor and Designer*. The Hague, 1991.
11. Louviere JJ, Hensher DA, Swait JD. *Stated Choice Methods. Analysis and Applications*. New York: Cambridge University Press, 2000.
12. Schwappach DLB, Koeck CM. Improving patient-provider communications: a web-access national health care survey panel. *Health Expect* 2003; **6**: 269–272.
13. Brick JM, Kalton G. Handling missing data in survey research. *Stat Methods Med Res* 1996; **5**: 215–238.
14. StataCorp. *Stata Statistical Software: Release 8.0*. College Station, TX: Stata Corporation, 2003.
15. Williams RL. A note on robust variance estimation for cluster-correlated data. *Biometrics* 2000; **56**: 645–646.
16. Long JS, Freese J. *Regression Models for Categorical Dependent Variables using STATA*. College Station, TX: Stata Corporation, 2001.
17. Kraman SS, Hamm G. Risk management: extreme honesty may be the best policy. *Ann Intern Med* 1999; **131**: 963–967.
18. Hebert PC, Levin AV, Robertson G. Bioethics for clinicians: 23. Disclosure of medical error. *CMAJ* 2001; **164**: 509–513.
19. Blendon RJ, DesRoches CM, Brodie M *et al*. Views of practicing physicians and the public on medical errors. *N Engl J Med* 2002; **347**: 1933–1940.
20. The Kaiser Family Foundation and Agency for Health Care Research and Quality (AHRQ). *National Survey on Americans as Health Care Consumers*. Washington, DC: AHRQ, 2000.
21. Blendon RJ, Schoen C, DesRoches C, Osborn R, Zapert K. Common concerns amid diverse systems: health care experiences in five countries. *Health Aff (Millwood)* 2003; **22**: 106–121.
22. National Patient Safety Foundation at the AMA. *Public Opinion of Patient Safety Issues. Research findings*. Washington, DC: National Patient Safety Foundation, 1997.
23. Forsa-Institut. *Behandlungsfehler. Repräsentative Befragung für die Techniker Krankenkasse (In German)*. Berlin: Techniker Krankenkasse, 2002.
24. Theobald A. *The World Wide Web as a Survey Tool (in German)*. Wiesbaden: Deutscher Universitäts-Verlag, 2000.
25. Levinsky NG. Truth or consequences. *N Engl J Med* 1998; **338**: 913–915.
26. Doyal L. The rationing debate. Rationing within the NHS should be explicit. The case for. *Br Med J* 1997; **314**: 1114–1118.
27. Hughes R. Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociol Health Illn* 1998; **20**: 381–400.
28. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician–patient communication. The relationship with malpractice claims among primary care physicians and surgeons. *JAMA* 1997; **277**: 553–559.
29. Finch J. The vignette technique in survey research. *Sociology* 1987; **21**: 105–114.
30. Liang BA. A system of medical error disclosure. *Qual Saf Health Care* 2002; **11**: 64–68.
31. Runciman WB, Merry AF, Tito F. Error, blame, and the law in health care—an antipodean perspective. *Ann Intern Med* 2003; **138**: 974–979.

Accepted for publication 14 March 2004

Appendix

AI: Vignette attributes, levels, and their phrasing

Attributes, variable name (abr.)	Attribute levels (vignette phrasing, translated from German original)
Severity of initial condition (sev.)	<p>Mild ('A patient with stomach ache and nausea undergoes a number of tests to determine the cause of his symptoms. In addition to blood and stool tests a gastroscopy is performed. Upon completion of all examinations the doctor prescribes a certain course of medication.')</p> <p>Moderate ('A patient had suffered from cough for a while. The doctor auscultated the patient's respiration and ordered blood tests. He reassured the patient that the cough would probably not have any serious origin. He referred the patient to obtain an X-ray of the lung and prescribed some drugs, which should alleviate breathing.')</p> <p>Severe ('A patient was treated in the intensive care unit for several days because of an acute myocardial infarction. At discharge, he still felt in a weak condition. A number of examinations and tests were taken and several drugs were prescribed. These should prevent a second infarction and improve the patient's well-being.')</p>
Type of error (type)	<p>Oversight of diagnostic information ('Even after a while, the patient still suffered from his symptoms. The doctor reviewed the medical record and discovered that one of the numerous test results had slipped back in the record and he had overlooked it. Due to this, he had misjudged and wrongly assessed the situation.')</p> <p>Oversight of allergy/contraindication ('The patient suffered an allergic reaction to the drugs. The allergy was known and documented in the medical record. The doctor, however, did not remember and overlooked the information when choosing the drugs.')</p> <p>Wrong treatment ('Following a worried phone call by the patient's family after the patient's condition severely worsened, the doctor reviews the patient's medical record and his prescribed treatment path. The doctor discovers that he prescribed a wrong combination of drugs to the patient.')</p>
Severity of error (severr.)	<p>Mild ('The incident had no serious consequences. But the success of treatment was delayed and the patient had to wait for unnecessarily long recovery.')</p> <p>Moderate ('The doctor immediately initiates the correct therapy. However, following the initial wrong treatment, the patient is confined to bed for more than 2 weeks, has a high temperature, needs additional care, and his recovery is rather slow. With accurate therapy from the beginning these negative consequences could most certainly have been avoided.')</p> <p>Severe ('The course of events required emergency care. During this, it was diagnosed that the patient had suffered a stroke. The patient has been impaired since then and has had to adapt his daily life completely. With accurate care the stroke could most certainly have been avoided.')</p>
Acknowledgment of error as such (ack.)	<p>Blank (no information provided)</p> <p>Yes ('Unfortunately, I have to tell you that I have made an error in your treatment.' Depends on type. For type = oversight of diagnostic information: 'I overlooked an important diagnostic information'. For type = oversight of allergy: 'I overlooked your allergy'. For type = wrong treatment: 'I prescribed a wrong combination of drugs'.)</p> <p>No ('Unfortunately, problems occurred in your treatment.' Depends on type. For type = oversight of diagnostic information: 'The origin of your symptoms was ambiguous'. For type = oversight of allergy: 'There were treatment side-effects'. For type = wrong treatment: 'You responded quite uncommonly to the treatment'.)</p>
Communication (com.)	<p>Depending on severr. and type</p> <p>Narrow:</p> <p>For type = oversight of diagnostic information: 'Therefore, a causal treatment was initially not possible'. For type = oversight of allergy: 'You probably had an allergic reaction to one of the drugs I prescribed'. For type = wrong treatment: 'These drugs were not effective in achieving the desired outcome'.</p> <p>Basic:</p> <p>As narrow, plus: for severr. = mild: 'Fortunately, this has no negative consequences for your health'. For severr. = moderate: 'As a result, your condition initially worsened, but fortunately we responded rapidly and rectified the situation. You can rest assured that there will be no long-term negative consequences for your health'. For severr. = severe: 'It is very likely, that this caused the stroke'.</p>

continued

A1: continued

	Comprehensive: As basic, plus: 'Let me assure you that I will investigate how this could have happened. We will do our best to avoid such events in the future. I will understand if you wish to discontinue your treatment with me, and can offer to refer you to a colleague of mine.'
Apology (apo.)	Blank (no information provided) Yes ('I'm very sorry that you had to go through this. Let me express my sincere apologies.')

A2: Example vignette (translated from German original)

Background information. A patient with stomach ache and nausea undergoes a number of tests to determine the cause of his symptoms. In addition to blood and stool tests, a gastroscopy is performed. Upon completion of all examinations the doctor prescribes a certain course of medication. Following a worried phone call by the patient's family after the patient's condition severely worsened, the doctor reviews the patient's medical record and his prescribed treatment path. The doctor discovers that he prescribed a wrong combination of drugs to the patient. The doctor immediately initiates the correct therapy. However, following the initial wrong treatment, the patient is confined to bed for more than 2 weeks, has a high temperature, needs additional care, and his recovery is rather

slow. With accurate therapy from the beginning these negative consequences could most certainly have been avoided.

In a personal conversation, the doctor talks to the patient: 'Unfortunately, I have to tell you that I have made an error in your treatment. I prescribed a wrong combination of drugs. These drugs were not effective in achieving the desired outcome. As a result, your condition initially worsened, but fortunately we responded rapidly and rectified the situation. You can rest assured that there will be no long-term negative consequences for your health, and that I will investigate how this could have happened. We will do our best to avoid such events in the future. I will understand if you wish to discontinue your treatment with me, and can offer to refer you to a colleague of mine. I'm very sorry that you had to go through this. Let me express my sincere apologies.'