

# Association of Perceived Medical Errors With Resident Distress and Empathy

## A Prospective Longitudinal Study

Colin P. West, MD, PhD

Mashele M. Huschka, BS

Paul J. Novotny, MS

Jeff A. Sloan, PhD

Joseph C. Kolars, MD

Thomas M. Habermann, MD

Tait D. Shanafelt, MD

**M**EDICAL ERRORS AND PATIENT safety are an important concern for patients and physicians. Medical errors have received increased attention since 1999, when the Institute of Medicine reported that up to 100 000 US patients die each year because of preventable adverse events.<sup>1</sup> The proportion of hospitalized patients affected by medical errors has been estimated to be 5% to 10%,<sup>2-6</sup> although it has approached 50% in some studies.<sup>7,8</sup> The morbidity, mortality, and financial costs of these events may be great.<sup>9-13</sup>

Many reports on medical errors have focused on the rate at which errors affect patients. Less commonly addressed is the proportion of physicians who commit errors. Several studies have evaluated this rate among resident physicians.<sup>14-16</sup> In a cross-sectional study examining self-defined errors, Mizrahi<sup>14</sup> found that 47% of internal medicine residents

See also pp 1049, 1055, 1063, and 1132.

**Context** Medical errors are associated with feelings of distress in physicians, but little is known about the magnitude and direction of these associations.

**Objective** To assess the frequency of self-perceived medical errors among resident physicians and to determine the association of self-perceived medical errors with resident quality of life, burnout, depression, and empathy using validated metrics.

**Design, Setting, and Participants** Prospective longitudinal cohort study of categorical and preliminary internal medicine residents at Mayo Clinic Rochester. Data were provided by 184 (84%) of 219 eligible residents. Participants began training in the 2003-2004, 2004-2005, and 2005-2006 academic years and completed surveys quarterly through May 2006. Surveys included self-assessment of medical errors and linear analog scale assessment of quality of life every 3 months, and the Maslach Burnout Inventory (depersonalization, emotional exhaustion, and personal accomplishment), Interpersonal Reactivity Index, and a validated depression screening tool every 6 months.

**Main Outcome Measures** Frequency of self-perceived medical errors was recorded. Associations of an error with quality of life, burnout, empathy, and symptoms of depression were determined using generalized estimating equations for repeated measures.

**Results** Thirty-four percent of participants reported making at least 1 major medical error during the study period. Making a medical error in the previous 3 months was reported by a mean of 14.7% of participants at each quarter. Self-perceived medical errors were associated with a subsequent decrease in quality of life ( $P=.02$ ) and worsened measures in all domains of burnout ( $P=.002$  for each). Self-perceived errors were associated with an odds ratio of screening positive for depression at the subsequent time point of 3.29 (95% confidence interval, 1.90-5.64). In addition, increased burnout in all domains and reduced empathy were associated with increased odds of self-perceived error in the following 3 months ( $P=.001$ ,  $P<.001$ , and  $P=.02$  for depersonalization, emotional exhaustion, and lower personal accomplishment, respectively;  $P=.02$  and  $P=.01$  for emotive and cognitive empathy, respectively).

**Conclusions** Self-perceived medical errors are common among internal medicine residents and are associated with substantial subsequent personal distress. Personal distress and decreased empathy are also associated with increased odds of future self-perceived errors, suggesting that perceived errors and distress may be related in a reciprocal cycle.

JAMA. 2006;296:1071-1078

www.jama.com

**Author Affiliations:** Division of General Internal Medicine, Department of Medicine (Dr West), Division of Biostatistics, Department of Health Sciences Research (Dr Sloan, Ms Huschka, and Mr Novotny), Division of Gastroenterology and Hepatology, Department of Medicine (Dr Kolars) and Division of Hematology, Department of Medicine (Drs

Habermann and Shanafelt), Mayo Clinic College of Medicine, Rochester, Minn.

**Corresponding Author:** Colin P. West, MD, PhD, Division of General Internal Medicine, Department of Medicine, Mayo Clinic College of Medicine, 200 First St SW, Rochester, MN 55905 (west.colin@mayo.edu).

reported making serious errors during their training. In another cross-sectional study of internal medicine residents, Wu et al<sup>15</sup> found that 45% of surveyed residents at various stages of training reported making at least 1 mistake (defined as “an act of omission by any caregiver which would have been judged wrong by knowledgeable peers at the time it occurred”) during their residency. More recently, Jagsi et al<sup>16</sup> surveyed residents across multiple specialties and found that 18% of respondents acknowledged at least 1 adverse event (defined as “a complication, injury, or harm to a patient resulting from medical management [not from the patient’s underlying condition or disease]”) in a patient under their care during the previous week. More than one third of these events was classified as a mistake by the resident, using the same definition as Wu et al.<sup>15</sup> An additional 23% of respondents reported a near miss (defined as “a mistake that does not reach the patient or if it reaches the patient does not result in injury or harm”) for which they felt at least partially responsible.

These studies suggest that a significant proportion of graduate medical trainees make medical errors. Committing errors can have a significant impact on clinicians. If patients are the first victims of medical errors, physicians have been termed the “second victims”<sup>17</sup> and commonly experience feelings of distress, guilt, shame, and depression in response to medical errors.<sup>15,18-23</sup> These effects may be long lasting,<sup>14,15,24-26</sup> with some physicians feeling “permanently wounded” as a result.<sup>20</sup>

To date, studies of the effect of medical errors on physicians have been descriptive cross-sectional evaluations that preclude quantitative assessment of the temporal relationship or magnitude of the association between errors and physician distress. We used a prospective longitudinal design (the Mayo Internal Medicine Well-being Study) to evaluate the frequency of perceived medical errors among internal medicine residents and to measure the association of these medical errors with

resident quality of life (QOL), burnout, symptoms of depression, and empathy using validated metrics.

## METHODS

### Participants

All entering categorical and preliminary internal medicine trainees in academic years 2003-2004, 2004-2005, and 2005-2006 at the Mayo Clinic Rochester Internal Medicine Residency program were eligible to participate. Residents in these academic classes had attended 108 US and international medical schools. These residents spend approximately half of their rotations on in-hospital services with overnight call responsibilities, with the remaining time spent in outpatient and subspecialty consultation rotations. Residents were invited to participate in this study during their orientation before beginning residency or by telephone if unable to attend orientation. Participation was elective. Individuals who desired to participate signed written informed consent to be surveyed every 3 months. The Mayo Clinic institutional review board approved this study.

### Data Collection

Residents were electronically surveyed every 3 months throughout their training. The current study focuses on data collected through May 2006. As of this date, categorical residents beginning training in 2003, 2004, and 2005 had received 12, 8, and 4 surveys, respectively. Preliminary (1-year) residents beginning training in 2003, 2004, and 2005 had received 5, 5, and 4 surveys, respectively. Surveys included questions about demographic characteristics, current rotation characteristics, coping strategies for dealing with stress, and report of self-perceived medical errors. Validated survey tools were used to measure QOL, burnout, symptoms of depression, and empathy, as described below. Self-reported medical errors and QOL were assessed quarterly; to avoid an excess burden on participants, burnout, symptoms of depression, and empathy were evaluated every 6 months. No mem-

ber of the Mayo Clinic Department of Medicine had access to identifying information on study items for individual participants. Nonidentifying numeric codes were used by statisticians to preserve resident anonymity when longitudinal data were collected and analyzed. Residents who screened positive for depression were identified by a separate study statistician and received a letter by certified mail notifying them of this result and informing them of confidential resources available to those desiring help.

### Study Measures

**Self-reported Medical Errors.** Perceived medical errors were evaluated by self-report every 3 months by asking residents, “Are you concerned you have made any major medical errors in the last 3 months?” The intent of this question was to identify events internalized by residents as major medical errors, rather than to document events associated with patient risk. Accordingly, major medical errors were not specifically defined for the residents. Thus, self-reported errors in this study represent major medical errors as perceived by each resident. Residents reporting errors were also asked to indicate to whom they had spoken about these errors.

**QOL, Burnout, and Depression.** Quality of life was measured by a single-item linear analog scale assessment (LASA). This instrument measures overall QOL on a 0 to 10 scale, with response anchors ranging from “as bad as it can be” (0) to “as good as it can be” (10). This scale has been validated across a wide range of medical conditions and populations.<sup>27-31</sup>

Burnout is a syndrome encompassing 3 domains (depersonalization, emotional exhaustion, and a sense of low personal accomplishment) that is associated with decreased work performance.<sup>32</sup> Burnout was measured using the Maslach Burnout Inventory (MBI), a 22-item tool evaluating each of these domains.<sup>32</sup> Responders rate the frequency with which they experience various feelings or emotions on a 7-point Likert scale, with response

options ranging from “never” to “daily.” Higher values of depersonalization (Maslach Burnout Inventory–Depersonalization) and emotional exhaustion (Maslach Burnout Inventory–Emotional Exhaustion [MBI-EE]) and lower values of personal accomplishment (Maslach Burnout Inventory–Personal Accomplishment) signify burnout. This instrument has been used in numerous previous studies of physicians.<sup>33-36</sup>

Depression screening used the 2-question approach described by Spitzer et al<sup>37</sup> and validated by Whooley et al.<sup>38</sup> This instrument has been used in a variety of patient populations,<sup>37,38</sup> including 2 studies of physicians.<sup>33,39</sup> This tool includes questions about depressed mood and anhedonia: “During the past month, have you often been bothered by feeling down, depressed, or hopeless?” and “During the past month, have you often been bothered by little interest or pleasure in doing things?” A positive screen for depression is defined as a “yes” response to either question. This screening instrument has reported positive likelihood ratios of up to 3.42 for the diagnosis of current major depression and negative likelihood ratios as low as 0.07.<sup>37,38</sup> These likelihood ratios are typical of other depression screening instruments reported in the literature.<sup>38,40</sup> Assuming a 25% prevalence of depression similar to that reported in other samples of internal medicine residents,<sup>33,36,41</sup> screening positive on this instrument implies a probability of depression of up to 53%, whereas a negative screen result implies a probability of depression as low as 2%.

**Empathy.** Empathy is a multidimensional construct with cognitive and emotive domains. Cognitive empathy relates to an individual’s ability to understand the perspective of another person about his or her circumstances. Emotive empathy refers to an individual’s concern for the feelings of others.<sup>42-46</sup> Empathy was measured with the Interpersonal Reactivity Index (IRI), a 28-item instrument with 4 separate 7-item subscales evaluating

different dimensions of empathy that are considered independently.<sup>42</sup> Each question item is scored on a Likert scale from 0 to 4, with response anchors ranging from “does not describe me well” (0) to “describes me well” (4), so that the maximum score for each subscale is 28. We included the IRI subscales measuring the cognitive and emotive domains of empathy according to previous studies demonstrating the utility of these subscales for evaluating empathy among resident physicians.<sup>36,46-48</sup>

### Statistical Analyses

Standard univariate statistics were used to characterize the sample. Comparisons between residents reporting errors and residents reporting no errors were initially made using summary statistics, collapsing responses within each individual into a single average outcome.<sup>49</sup> These comparisons were analyzed using the Wilcoxon-Mann-Whitney test for continuous variables and the Fisher exact test for proportions.

To incorporate the repeated-measures study design, the association of self-perceived errors with QOL, empathy, burnout, and depression was evaluated using generalized estimating equations (GEE), an extension of generalized linear models that accounts for correlated repeated measurements within individuals.<sup>49,50</sup> Analyses were performed examining the association of self-perceived errors with distress at the subsequent time point. Because the surveys asked about self-perceived errors during the previous 3 months, these errors preceded the assessment of all distress variables except depression, for which some overlap occurred because the depression screening tool assesses symptoms of depression during the previous 4 weeks.

Finally, we analyzed the association of distress and empathy with the likelihood of a self-perceived error during the following 3 months. For these analyses, the assessment of all distress variables preceded the self-reported errors.

Because of multicollinearity among distress variables, each model included self-reported errors and 1 distress variable. To properly calculate variance terms for repeated-measures analyses, the GEE method requires that a correlation structure be specified. Selecting the correct correlation structure for GEE analyses does not in general affect parameter estimation but does allow more precise estimates. Where allowed by the data, we specified unstructured correlations. Alternative structures in order of preference were autoregressive and exchangeable.

Statistical analyses were conducted with SAS version 8.2 (SAS Institute Inc, Cary, NC). Statistical significance was set at the .05 level, and all tests were 2-tailed.

## RESULTS

Participants were 184 (84%) of 219 eligible residents; there were no statistically significant differences in age, sex, or program type between participants and nonparticipants. The demographic characteristics of study participants are shown in TABLE 1. Age 30 years was used as a threshold to approximately separate residents with more standard medical education histories from those who may have had other life experiences before beginning their residency. The categories for debt were intended to be comparable with those of a study that investigated the relationship between debt level and stress.<sup>51</sup> Of the participants, 100% completed at least 1 survey during the study period, with response rates to individual surveys ranging from 64% to 94% (mean, 72.2%). Baseline participant characteristics for QOL, burnout, depression screening, and empathy are shown in TABLE 2.

Errors were reported in 130 (14.7%) of 883 resident-quarters. Perceived error rates by quarter of training ranged from 4.3% to 23.1%. Overall, 34% of study participants reported at least 1 major medical error during the study period, and 43% of residents completing at least 1 year of training reported errors. Of the participants, 20% re-

ported 1 error, 6% reported 2 errors, and 8% reported 3 or more errors during the study period. Self-perceived error rates did not vary significantly by age, sex, program type, amount of student loan debt, relationship status, or parental status. Of perceived errors, 97% were discussed with at least 1 individual. The most common group with whom errors were discussed was other

residents (83%), although a majority of residents also reported discussing perceived errors with close family and friends (65%) or supervisory faculty (54%).

Summary measures to identify general associations between self-perceived errors and resident QOL, burnout, symptoms of depression, and empathy are shown in Table 2. Residents reporting at least 1 error during the study period had significantly lower overall QOL on the LASA ( $-0.52$ ;  $P = .03$ ). Residents reporting errors also had higher levels of burnout, as evidenced by increased depersonalization ( $+3.23$ ;  $P < .001$ ), increased emotional exhaustion ( $+6.85$ ;  $P < .001$ ), and a lower sense of personal accomplishment ( $-2.99$ ;  $P = .001$ ) on the MBI. More than 60% of residents reporting an error screened positive for depression at least once during the study period, nearly twice the rate in residents reporting no errors. Residents reporting errors had nonstatistically significant lower emotive empathy scores ( $-0.89$ ;  $P = .15$ ) and cognitive empathy scores ( $-0.65$ ;  $P = .31$ ) on the IRI.

The association of a self-perceived error with QOL, burnout, symptoms of depression, and empathy at the subsequent survey time point is shown in

TABLE 3. Self-perceived medical errors had a statistically significant adverse association with overall QOL, all 3 domains of burnout, and the likelihood of screening positive for depression. For example, a self-perceived major medical error was associated with a 4.58-point increase in emotional exhaustion on the MBI-EE scale and with an increased odds of 3.29 (95% confidence interval, 1.90-5.64) for a positive depression screen result at the subsequent survey time point.

The association between distress at each survey point and a self-perceived error in the subsequent 3 months is shown in TABLE 4. Diminished empathy and higher levels of burnout in all domains were associated with increased odds of a self-perceived error in the subsequent 3 months. Each 1-point increase in depersonalization and emotional exhaustion score was associated with a 10% and 7% increase, respectively, in the odds of reporting an error in the following 3 months. Similarly, each 1-point increase in personal accomplishment, emotive empathy, and cognitive empathy score was associated with a 7%, 9%, and 9% decrease, respectively, in the odds of a self-perceived error in the following 3 months.

**Table 1.** Demographic Characteristics of Participants at Entry to Study (N = 184)

Variable Category	No. (%)
Age, y	
≤30	129 (70.1)
>30	30 (16.3)
Missing	25 (13.6)
Sex	
Male	94 (51.1)
Female	66 (35.9)
Missing	24 (13.0)
Program	
Categorical	126 (68.5)
Preliminary	58 (31.5)
Student loan debt, \$	
<50 000	68 (37.0)
50 000-100 000	25 (13.6)
>100 000	67 (36.4)
Missing	24 (13.0)
Relationship status	
Single	67 (36.4)
Married	81 (44.0)
Divorced	5 (2.7)
Partner	7 (3.8)
Missing	24 (13.0)
Children	
Yes	27 (14.7)
No	133 (72.3)
Missing	24 (13.0)

**Table 2.** Quality of Life, Burnout, Symptoms of Depression, and Empathy Measures for Residents Reporting No Perceived Errors vs Reporting Perceived Errors\*

Variable	Metric (Scale)	Group Baseline, Mean (SD) (N = 184)	No Reported Errors (n = 122)	Reported Errors (n = 62)	Difference (95% Confidence Interval)	P Value
QOL	LASA overall QOL (0-10), mean	6.60 (1.88) (n = 160)	6.54	6.01	-0.52 (-1.00 to -0.05)	.03†
Burnout‡						
Depersonalization	MBI-DP (0-30), mean	7.10 (5.94) (n = 145)	6.62	9.85	3.23 (1.35 to 5.12)	<.001†
Emotional exhaustion	MBI-EE (0-54), mean	21.51 (9.91) (n = 142)	19.21	26.06	6.85 (3.88 to 9.82)	<.001†
Personal accomplishment	MBI-PA (0-48), mean	39.01 (5.25) (n = 142)	39.26	36.27	-2.99 (-4.77 to -1.22)	.001†
Depression	Any positive 2-item depression screen, %	32.21 (46.99) (n = 149)	33.02	63.33	3.50 (1.71 to 7.20)§	<.001
Empathy						
Emotive	IRI-EC (0-28), mean	22.47 (4.26) (n = 159)	22.25	21.36	-0.89 (-2.11 to 0.32)	.15†
Cognitive	IRI-PT (0-28), mean	20.25 (4.48) (n = 158)	20.60	19.95	-0.65 (-1.91 to 0.60)	.31†

Abbreviations: IRI-EC, Interpersonal Reactivity Index–Empathic Concern Subscale; IRI-PT, Interpersonal Reactivity Index–Perspective Taking Subscale; LASA, linear analog scale assessment; MBI-DP, Maslach Burnout Inventory–Depersonalization; MBI-EE, Maslach Burnout Inventory–Emotional Exhaustion; MBI-PA, Maslach Burnout Inventory–Personal Accomplishment; QOL, quality of life.

\*Summary statistics averaged over all survey points providing data.

†Wilcoxon-Mann-Whitney test.

‡Higher depersonalization or emotional exhaustion scores and lower personal accomplishment scores are indicative of greater burnout. Thresholds to categorize physicians as having low, average, or high burnout are based on normative scales<sup>32</sup> (depersonalization: low burnout, 0 to 5; average burnout, 6 to 9; high burnout, ≥10; emotional exhaustion: low burnout, 0 to 18; average burnout, 19 to 26; high burnout, ≥27; personal accomplishment: low burnout, ≥40; average burnout, 34 to 39; high burnout, 0 to 33).

§Odds ratio for a positive depression screen for the errors group relative to the no-errors group.

||Fisher exact test.

The addition of other potential confounding or interacting factors (type of clinical rotation, self-reported satisfaction with work-life balance, occurrence of a major negative life event [eg, divorce, death in the family], occurrence of a major positive life event [eg, marriage, birth in the family], and preferred coping strategies) to these models did not significantly alter the results. Residents in training for a shorter duration necessarily contributed fewer time points to this study. To evaluate this potential source of bias, analyses examining only residents who had completed their first year of training were conducted and yielded similar results.

## COMMENT

This prospective longitudinal study shows that self-perceived major medical errors were common among the studied internal medicine residents, with about one third of participants reporting a major error at least once during the study period. In addition to their potential effects on patients, perceived medical errors exhibited a strong association with multiple domains of physicians' personal well-being. In particular, self-perceived errors were associated with decreased QOL and increases in burnout and symptoms of depression. Declines in empathy were also associated with perceived medical errors.

Previous research has demonstrated that changes in LASA scores approaching 0.5 SDs or more, as we observed, are clinically significant.<sup>52</sup> Additionally, our reported associations of perceived errors with burnout are large, given that "low" and "high" categories on the MBI scales are separated by only 5 points (depersonalization), 9 points (emotional exhaustion), and 7 points (personal accomplishment).<sup>32</sup> Bellini et al<sup>47</sup> suggested that 1.0- to 1.5-point changes in resident empathy scores on the IRI are important, although it remains unclear what constitutes a clinically meaningful change in empathy score. Overall, this evidence suggests that our findings are not only statistically significant but also clinically meaningful.

When considered with previous studies demonstrating a link between personal distress and empathy,<sup>36,46</sup> these results imply that medical errors represent an important contributor to the personal distress and loss of compassion reported in numerous studies of residents,<sup>36,47,48</sup> which is significant not only because the personal effects of making an error can be profound but also because personal distress appears to negatively affect patient care.<sup>33,34,53,54</sup> Our results also support the link between physician distress and subsequent self-

reported errors. Taken together, these results suggest a vicious cycle whereby medical errors may lead to personal distress, which then contributes to further deficits in patient care.

In response to these findings, one step for residency programs should be to reduce medical errors to the greatest extent possible. System issues are cited as major contributors to medical error,<sup>1,55-58</sup> and system-based solutions have been explored,<sup>59,60</sup> including interventions designed to reduce medication errors, control nosoco-

**Table 3.** Association of a Self-Perceived Major Medical Error in the Previous 3 Months With Quality of Life, Burnout, Symptoms of Depression, and Empathy (N = 184)

Dependent Variable	Metric (Scale)	Parameter Estimate (95% Confidence Interval)*	P Value†
QOL	LASA overall QOL (0-10)	-0.39 (-0.72 to -0.06)	.02
Burnout‡			
Depersonalization	MBI-DP (0-30)	2.45 (0.94 to 3.97)	.002
Emotional exhaustion	MBI-EE (0-54)	4.58 (1.71 to 7.46)	.002
Personal accomplishment	MBI-PA (0-48)	-2.59 (-4.22 to -0.97)	.002
Depression	Any positive 2-item depression screen	3.29 (1.90 to 5.64)§	<.001
Empathy			
Emotive	IRI-EC (0-28)	-0.56 (-1.39 to 0.28)	.19
Cognitive	IRI-PT (0-28)	-0.72 (-1.59 to 0.15)	.10

Abbreviations: IRI-EC, Interpersonal Reactivity Index–Empathic Concern Subscale; IRI-PT, Interpersonal Reactivity Index–Perspective Taking Subscale; LASA, linear analog scale assessment; MBI-DP, Maslach Burnout Inventory–Depersonalization; MBI-EE, Maslach Burnout Inventory–Emotional Exhaustion; MBI-PA, Maslach Burnout Inventory–Personal Accomplishment; QOL, quality of life.

\*Errors coded as 0/1 (no/yes). Parameter estimates indicate the change in each metric associated with a self-reported error. For example, a self-reported error in the previous 3 months is associated with a 2.45-unit increase in the depersonalization score.

†Using generalized estimating equation models adjusted for time. Working correlations were unstructured where possible and autoregressive or exchangeable otherwise as allowed by data.

‡See footnote to Table 2 for scale classifications.

§Odds ratio for a positive depression screen result, given a perceived error.

**Table 4.** Association of Quality of Life, Burnout, Symptoms of Depression, and Empathy With a Self-Perceived Major Medical Error in the Following 3 Months (N = 184)

Independent Variable	Metric (Scale)	Odds Ratio (95% Confidence Interval)*	P Value†
QOL	LASA overall QOL (0-10)	0.93 (0.83-1.04)	.20
Burnout‡			
Depersonalization	MBI-DP (0-30)	1.10 (1.04-1.16)	.001
Emotional exhaustion	MBI-EE (0-54)	1.07 (1.03-1.12)	<.001
Personal accomplishment	MBI-PA (0-48)	0.93 (0.88-0.99)	.02
Depression	Any positive 2-item depression screen	1.93 (0.93-3.99)	.08
Empathy			
Emotive	IRI-EC (0-28)	0.91 (0.84-0.98)	.02
Cognitive	IRI-PT (0-28)	0.91 (0.85-0.98)	.01

Abbreviations: IRI-EC, Interpersonal Reactivity Index–Empathic Concern Subscale; IRI-PT, Interpersonal Reactivity Index–Perspective Taking Subscale; LASA, linear analog scale assessment; MBI-DP, Maslach Burnout Inventory–Depersonalization; MBI-EE, Maslach Burnout Inventory–Emotional Exhaustion; MBI-PA, Maslach Burnout Inventory–Personal Accomplishment; QOL, quality of life.

\*Odds ratio of a self-reported error in the following 3 months associated with a 1-unit increase in each distress metric. †Using generalized estimating equation models adjusted for time. Working correlations were unstructured where possible and autoregressive or exchangeable otherwise as allowed by data.

‡See footnote to Table 2 for scale classifications.

mial infection rates, and prevent falls among hospitalized patients. Within graduate medical education, training environments that result in excessive resident fatigue have been targeted by duty-hour reforms.<sup>61-63</sup> However, all components of the Accreditation Council for Graduate Medical Education resident duty-hour limitations<sup>62</sup> were in effect, with documented compliance, at our training program throughout the entire study, suggesting that the self-reported error rates we observed already reflect the benefits of this intervention.

System efforts are unlikely to eliminate errors completely, and when errors do occur, physicians often have limited resources on which to call for support.<sup>17,64,65</sup> In our study, the majority of residents discussed their perceived errors with colleagues, supervisory staff, and close family or friends. Additional coping strategies reported in the literature include discussing errors with patients, accepting responsibility, and working to put error prevention methods in place after error analysis.<sup>18,23,64,66</sup> Specific curricula on personal awareness and self-care to promote strategies for coping with the emotional impact of errors are needed but have been slow to develop.<sup>20,67-69</sup> Additional research is required to identify effective approaches to assist physicians who have made medical errors.

Our results also suggest that residency programs should ensure that efforts are in place to prevent, identify, and treat burnout and to promote empathy and well-being for the welfare of residents and patients. Further investigations of the impact of personal distress on error occurrence rates will help clarify the roles that issues such as physician QOL, burnout, and depression play in patient safety.

Our study has several limitations. First, the definition and reporting of major medical errors in this study were based on self-perception. The extent to which these self-perceived errors accurately reflect the frequency of medical errors and whether these perceived errors actually affected patient

outcomes cannot be determined. Although no single method of measuring errors is ideally suited to all purposes,<sup>70</sup> this approach reflects the study's aim of measuring the effects of perceived errors on physician well-being: errors physicians are not aware of would not be expected to have any such effects. Previous work has suggested that physician-identified adverse events differ from events identified by medical record review but that such self-reported adverse events may be more likely to represent preventable medical errors.<sup>71</sup> It is also possible that a perceived error that results in adverse patient consequences could affect the level of distress experienced by the resident. Because we did not assess patient consequences, we are unable to address this issue.

Second, the generalizability of these results from a single academic medical center to other training programs is unknown. Our participation and survey response rates were high relative to those of other physician surveys,<sup>72,73</sup> suggesting that the results of this study are representative of residents in our training program. Residents in this program attended a wide range of medical schools and work in clinical settings characteristic of academic residency training programs, suggesting they are likely to be representative of other academic training programs. In addition, the error rates,<sup>14-16</sup> burnout scores,<sup>33-36,39</sup> rates of a positive depression screen result,<sup>33,39</sup> and empathy scores<sup>36,47,48</sup> observed in this study were similar to those found in other samples of medical residents.

Third, our survey structure measured distress at defined survey points, but unmeasured distress could occur between survey time points and therefore could precede a perceived error within each period. Additionally, it is possible that retrospective error reporting is distorted by feelings of distress. Although it is unclear whether such feelings of distress would make reporting of errors more likely or less likely, the potential exists for current distress to influence error reporting and

therefore affect our results. Because of these issues, our results are best interpreted as associations rather than as definitive evidence of causation.

Fourth, the depression screening instrument we used cannot diagnose depression by itself. Although the positive likelihood ratio for this instrument is similar to that of other accepted depression screening tools,<sup>38,40</sup> the post-test probability of about 50% described previously means that additional evaluation would be necessary to diagnose depression in residents with positive screen results. Our findings suggest that a positive depression screening result is associated with self-perceived errors, but because of these limitations, further study is needed to conclusively link clinical depression with medical errors.

Fifth, some potential confounding variables could not be evaluated. For example, it is possible that personality traits such as being highly self-critical, confident, or reflective affect some aspects of how physicians perceive or respond to errors. Also, sleep deprivation could contribute to medical errors and resident distress and was not assessed in this study beyond documented institutional compliance with resident duty-hour limitations.

Finally, the models reported in this study evaluate the relationship between self-perceived errors and each distress variable individually. Because of multicollinearity between the distress variables, we have limited ability to separate the effects of individual distress variables from one another.

In summary, these results suggest that self-perceived medical errors are common among internal medicine residents. These errors are associated with significant subsequent personal distress. Burnout and loss of empathy are also associated with an increased risk of future self-perceived major medical errors. A majority of residents discuss their errors with colleagues, supervising faculty, or friends and family, but formal programs to provide additional support for physicians who make errors appear warranted. Further in-

vestigation to identify the most effective posterror support mechanisms is needed in parallel with ongoing system efforts to reduce error rates and resident distress.

**Author Contributions:** Dr West had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** West, Sloan, Kolars, Habermann, Shanafelt.

**Acquisition of data:** West, Huschka, Novotny, Sloan, Kolars, Habermann, Shanafelt.

**Analysis and interpretation of data:** West, Huschka, Novotny, Sloan, Shanafelt.

**Drafting of the manuscript:** West, Shanafelt.

**Critical revision of the manuscript for important intellectual content:** West, Huschka, Novotny, Sloan, Kolars, Habermann, Shanafelt.

**Statistical analysis:** West, Huschka, Novotny, Sloan.

**Obtained funding:** Sloan, Kolars, Habermann, Shanafelt.

**Administrative, technical, or material support:** Kolars, Habermann.

**Study supervision:** Sloan, Kolars, Habermann, Shanafelt.

**Financial Disclosures:** None reported.

**Funding/Support:** This work was supported by a Medicine Innovation Development and Advancement System grant from the Mayo Clinic Department of Medicine.

**Role of the Sponsor:** The funding source for this study played no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in the preparation of the manuscript. The funding source did review and approve the manuscript.

## REFERENCES

- Kohn LT, Corrigan JM, Donaldson MS, eds. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press; 1999.
- Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study. *N Engl J Med*. 1991;324:370-376.
- Gawande AA, Thomas EJ, Zinner MJ, Brennan TA. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery*. 1999;126:66-75.
- Chaudhry SI, Olofinboba KA, Krumholz HM. Detection of errors by attending physicians on a general medicine service. *J Gen Intern Med*. 2003;18:595-600.
- Baker GR, Norton PG, Flintoff V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ*. 2004;170:1678-1686.
- Graf J, von den Driesch A, Koch KC, Janssens U. Identification and characterization of errors and incidents in a medical intensive care unit. *Acta Anaesthesiol Scand*. 2005;49:930-939.
- Steel K, Gertman PM, Crescenzi C, Anderson J. Iatrogenic illness on a general medical service at a university hospital. *N Engl J Med*. 1981;304:638-642.
- Andrews LB, Stocking C, Krizek T, et al. An alternative strategy for studying adverse events in medical care. *Lancet*. 1997;349:309-313.
- Kalish RL, Daley J, Duncan CC, Davis RB, Coffman GA, Iezzoni LI. Costs of potential complications of care for major surgery patients. *Am J Med Qual*. 1995;10:48-54.
- Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients: excess length of stay, extra costs, and attributable mortality. *JAMA*. 1997;277:301-306.
- Bates DW, Spell N, Cullen DJ, et al. The costs of adverse drug events in hospitalized patients. *JAMA*. 1997;277:307-311.
- Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA*. 2003;290:1868-1874.
- Rothschild JM, Landrigan CP, Cronin JW, et al. The Critical Care Safety Study: the incidence and nature of adverse events and serious medical errors in intensive care. *Crit Care Med*. 2005;33:1694-1700.
- Mizrahi T. Managing medical mistakes: ideology, insularity, and accountability among internists-in-training. *Soc Sci Med*. 1984;19:135-146.
- Wu AW, Folkman S, McPhee SJ, Lo B. Do house officers learn from their mistakes? *JAMA*. 1991;265:2089-2094.
- Jagsi R, Kitch BT, Weinstein DF, Campbell EG, Hutter M, Weissman JS. Residents report on adverse events and their causes. *Arch Intern Med*. 2005;165:2607-2613.
- Wu AW. Medical error: the second victim. *BMJ*. 2000;320:726-727.
- Newman MC. The emotional impact of mistakes on family physicians. *Arch Fam Med*. 1996;5:71-75.
- Penson RT, Svendsen SS, Chabner BA, Lynch TJ Jr, Levinson W. Medical mistakes: a workshop on personal perspectives. *Oncologist*. 2001;6:92-99.
- Wears RL, Wu AW. Dealing with failure: the aftermath of errors and adverse events. *Ann Emerg Med*. 2002;39:344-346.
- Vincent C. Understanding and responding to adverse events. *N Engl J Med*. 2003;348:1051-1056.
- Hobgood C, Hevia A, Tamayo-Sarver JH, Weiner B, Riviello R. The influence of the causes and contexts of medical errors on emergency medicine residents' responses to their errors: an exploration. *Acad Med*. 2005;80:758-764.
- Engel KG, Rosenthal M, Sutcliffe KM. Residents' responses to medical error: coping, learning, and change. *Acad Med*. 2006;81:86-93.
- Christensen JF, Levinson W, Dunn PM. The heart of darkness: the impact of perceived mistakes on physicians. *J Gen Intern Med*. 1992;7:424-431.
- Cunningham W. The immediate and long-term impact on New Zealand doctors who receive patient complaints. *N Z Med J*. 2004;117:U972.
- Rowe M. Doctors' responses to medical errors. *Crit Rev Oncol Hematol*. 2004;52:143-145.
- Spitzer WO, Dobson AJ, Hall J, et al. Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. *J Chronic Dis*. 1981;34:585-597.
- Grunberg SM, Groshen S, Steingass S, Zaretsky S, Meyerowitz B. Comparison of conditional quality of life terminology and visual analogue scale measurements. *Qual Life Res*. 1996;5:65-72.
- Gudex C, Dolan P, Kind P, Williams A. Health state valuations from the general public using the visual analogue scale. *Qual Life Res*. 1996;5:521-531.
- Shanafelt TD, Novotny P, Johnson ME, et al. The well-being and personal wellness promotion strategies of medical oncologists in the North Central Cancer Treatment Group. *Oncology*. 2005;68:23-32.
- Rumman TA, Clark MM, Sloan JA, et al. Impacting quality of life for patients with advanced cancer with a structured multidisciplinary intervention: a randomized controlled trial. *J Clin Oncol*. 2006;24:635-642.
- Maslach C, Jackson SE, Leiter MP. *Maslach Burnout Inventory Manual*. 3rd ed. Palo Alto, Calif: Consulting Psychologists Press; 1996.
- Shanafelt TD, Bradley KA, Wipf JE, Back AL. Burnout and self-reported patient care in an internal medicine residency program. *Ann Intern Med*. 2002;136:358-367.
- Thomas NK. Resident burnout. *JAMA*. 2004;292:2880-2889.
- Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV. Burnout and internal medicine resident work-hour restrictions. *Arch Intern Med*. 2005;165:2595-2600.
- Rosen IM, Gimotty PA, Shea JA, Bellini LM. Evolution of sleep quantity, sleep deprivation, mood disturbances, empathy, and burnout among interns. *Acad Med*. 2006;81:82-85.
- Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 Study. *JAMA*. 1994;272:1749-1756.
- Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: two questions are as good as many. *J Gen Intern Med*. 1997;12:439-445.
- Goitein L, Shanafelt TD, Wipf JE, Slatore CG, Back AL. The effects of work-hour limitations on resident well-being, patient care, and education in an internal medicine residency program. *Arch Intern Med*. 2005;165:2601-2606.
- Williams JW Jr, Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? *JAMA*. 2002;287:1160-1170.
- Reuben DB. Depressive symptoms in medical house officers: effects of level of training and work rotation. *Arch Intern Med*. 1985;145:286-288.
- Davis M. A multidimensional approach to individual differences in empathy. *JSAS Catalog Selected Documents Psychol*. 1980;10:85.
- Davis M. Measuring individual differences in empathy: evidence for a multi-dimensional approach. *J Pers Soc Psychol*. 1983;44:113-126.
- Halpern J. What is clinical empathy? *J Gen Intern Med*. 2003;18:670-674.
- Hojat M, Gonnella JS, Nasca TJ, Mangione S, Vergare M, Magee M. Physician empathy: definition, components, measurement, and relationship to gender and specialty. *Am J Psychiatry*. 2002;159:1563-1569.
- Shanafelt TD, West C, Zhao X, et al. Relationship between increased personal well-being and enhanced empathy among internal medicine residents. *J Gen Intern Med*. 2005;20:559-564.
- Bellini LM, Baime M, Shea JA. Variation of mood and empathy during internship. *JAMA*. 2002;287:3143-3146.
- Bellini LM, Shea JA. Mood change and empathy decline persist during three years of internal medicine training. *Acad Med*. 2005;80:164-167.
- Diggle PJ, Liang KY, Zeger SL. *Analysis of Longitudinal Data*. Oxford, England: Clarendon Press; 1994.
- Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13-22.
- Collier VU, McCue JD, Markus A, Smith L. Stress in medical residency: status quo after a decade of reform? *Ann Intern Med*. 2002;136:384-390.
- Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. *J Clin Epidemiol*. 2005;58:1217-1219.
- Firth-Cozens J, Greenhalgh J. Doctors' perceptions of the links between stress and lowered clinical care. *Soc Sci Med*. 1997;44:1017-1022.
- Haas JS, Cook EF, Puopolo AL, Burstin HR, Cleary PD, Brennan TA. Is the professional satisfaction of general internists associated with patient satisfaction? *J Gen Intern Med*. 2000;15:122-128.
- Leape LL. Error in medicine. *JAMA*. 1994;272:1851-1857.
- Leape LL, Bates DW, Cullen DJ, et al; ADE Prevention Study Group. Systems analysis of adverse drug events. *JAMA*. 1995;274:35-43.
- Leape LL, Berwick DM. Five years after *To Err Is Human*: what have we learned? *JAMA*. 2005;293:2384-2390.
- Longo DR, Hewett JE, Ge B, Schubert S. The long road to patient safety: a status report on patient safety systems. *JAMA*. 2005;294:2858-2865.

59. Shojania KG, Duncan BW, McDonald KM, Wachter RM, eds. *Making Health Care Safer: A Critical Analysis of Patient Safety Practices*. Rockville, Md: Agency for Healthcare Research and Quality; 2001. Evidence Report/Technology Assessment No. 43; AHRQ publication 01-E058.
60. Ioannidis JP, Lau J. Evidence on interventions to reduce medical errors: an overview and recommendations for future research. *J Gen Intern Med*. 2001;16:325-334.
61. Gaba DM, Howard SK. Patient safety: fatigue among clinicians and the safety of patients. *N Engl J Med*. 2002;347:1249-1255.
62. Philibert I, Friedmann P, Williams WT; ACGME Work Group on Resident Duty Hours. New requirements for resident duty hours. *JAMA*. 2002;288:1112-1114.
63. Landrigan CP, Rothschild JM, Cronin JW, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med*. 2004;351:1838-1848.
64. Wu AW, Folkman S, McPhee SJ, Lo B. How house officers cope with their mistakes. *West J Med*. 1993;159:565-569.
65. Goldberg RM, Kuhn G, Andrew LB, Thomas HA Jr. Coping with medical mistakes and errors in judgment. *Ann Emerg Med*. 2002;39:287-292.
66. Lester H, Tritter JQ. Medical error: a discussion of the medical construction of error and suggestions for reforms of medical education to decrease error. *Med Educ*. 2001;35:855-861.
67. Novack DH, Suchman AL, Clark W, et al; Working Group on Promoting Physician Personal Awareness, American Academy on Physician and Patient. Calibrating the physician: personal awareness and effective patient care. *JAMA*. 1997;278:502-509.
68. Epstein RM. Mindful practice. *JAMA*. 1999;282:833-839.
69. Luszczak M, Kharod CU, Abbrescia K. Medical mistakes, errors in judgment, and personal awareness. *Ann Emerg Med*. 2003;41:155-156.
70. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med*. 2003;18:61-67.
71. O'Neil AC, Petersen LA, Cook EF, Bates DW, Lee TH, Brennan TA. Physician reporting compared with medical-record review to identify adverse medical events. *Ann Intern Med*. 1993;119:370-376.
72. Asch DA, Jedrzejewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol*. 1997;50:1129-1136.
73. Kellerman SE, Herold J. Physician response to surveys: a review of the literature. *Am J Prev Med*. 2001;20:61-67.

For my judgment is that it is much better that you should learn the manner of cutting by eye and touch than by reading and listening. For reading alone never taught anyone how to sail a ship, to lead an army, nor to compound a medicine, which is done rather by the use of one's own sight and the training of one's own hands.

—Jacobus Sylvius (1478-1555)