Hindsight Bias Among Psychiatrists

H. W. LeBourgeois III, MD, Debra A. Pinals, MD, Valerie Williams, MA, MS, and Paul S. Appelbaum, MD

It is crucial to minimize bias when offering forensic opinions; however, to our knowledge there are few, if any, existing data examining whether psychiatrists are susceptible to one source of such bias, hindsight bias. In the current study, 235 general and forensic psychiatrists reviewed hypothetical cases in which patients with suicidal or homicidal ideation presented for psychiatric care. We informed half of the participants that a suicide or homicide had occurred shortly after the patients were released from care (hindsight group) but withheld outcome information from the other participants (control group). Participants estimated the likelihood that suicide or violence would occur at the time of the patient's release and whether the standard of care had been met in each case. Responses were compared between groups for suggestions of hindsight bias. Results indicate that hindsight bias plays a role in assessments of risk, but not of negligence, and that psychiatrists who are American Academy of Psychiatry and the Law (AAPL) members may be less prone to respond with hindsight bias than are others.

J Am Acad Psychiatry Law 35:67–73, 2007

Psychiatrists, particularly forensic psychiatrists, may be asked to review cases in which a psychiatrist provided care to a patient in which there had been an adverse outcome such as suicide or violence. They are asked to provide an opinion about whether the treating psychiatrist properly assessed the risk of suicide or violence and met the standard of care in managing the patient's risk and in providing treatment. Psychiatrists who participate in case reviews do so in a variety of contexts, such as participation in medical review panels or serving as potential expert witnesses for an attorney or in administrative capacities that entail monitoring the quality of care provided by staff members. If the reviewing psychiatrist is of the opinion that the treating psychiatrist did not perform a proper risk assessment or did not meet the standard of care for managing risk, or both, the treating psychiatrist may be subject to disciplinary action or civil liability (e.g., malpractice suits). Therefore, it is crucial for those performing case reviews to minimize potential sources of bias when offering such opinions. This stance is consistent with the Ethics Guidelines for the Practice of Forensic Psychiatry put forth by the American Academy of Psychiatry and the Law (AAPL)[1] and with several publications that encourage those engaging in forensic work to strive for objectivity.[2–4]

Because such case reviews are necessarily performed after the fact, one source of bias may be hindsight bias.[5] Hindsight bias, also known as outcome bias,[6] is the tendency for persons equipped with knowledge of an outcome to exaggerate their ability to predict the inevitability of the outcome.[7] One who is subject to hindsight bias may simplify, trivialize, and retrospectively criticize the decisions of the treating doctor[8] without appreciating the contemporaneous difficulty of the decisions involved.

Findings in studies have demonstrated that physicians are susceptible to hindsight bias.[9–11] In one of these studies, anesthesiologists experienced in case review were provided sets of clinical case scenarios with the same descriptive facts, but with adverse anesthetic outcomes randomly causing either temporary or permanent injury. The study revealed that anesthesiologists were more likely to rate care as substandard in cases in which they were provided data showing that the patient had sustained permanent injury, despite the cases' having identical descriptive

Dr. LeBourgeois is Director, Forensic Psychiatry Fellowship Program and Assistant Professor of Psychiatry, Department of Psychiatry and Neurology, Tulane University School of Medicine, New Orleans, LA. Dr. Pinals is Director, Forensic Psychiatry Fellowship Program and Associate Professor of Psychiatry, Law and Psychiatry Program, and Ms. Williams is a Research Instructor, Center for Mental Health Services Research, Department of Psychiatry, University of Massachusetts Medical School, Worcester, MA. Dr. Appelbaum is Professor and Director, Division of Psychiatry, Law and Ethics, Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY. Presented at the 2004 Annual Meeting of the American Academy of Psychiatry and the Law in Scottsdale, Arizona, as a Research-in-Progress presentation. Address correspondence to: H. W. LeBourgeois III, MD, Department of Psychiatry and Neurology, Tulane University School of Medicine, 1440 Canal Street, TB-53, 10th Floor, New Orleans, LA 70112. E-mail: hlebour@tulane.edu

facts. The study's authors concluded that knowledge of outcome influences anesthesiologists' retrospective judgments of appropriateness of care delivered by other physicians.[11]

Although the question of whether psychiatrists are susceptible to hindsight bias has not been studied to our knowledge, previous research by LaBine and LaBine[5] has demonstrated that mock jurors reviewing therapists' actions in *Tarasoff*-type cases are susceptible to hindsight bias. In this study, a random sample of community residents was asked to read clinical case scenarios in which a therapist treated a patient with one of three outcomes: the patient became violent, the patient did not become violent, or the outcome of the case was not specified. The results showed that respondents who were informed that the patient had become violent were more likely to rate the therapist's actions as less reasonable. In their discussion of the study, the authors questioned whether mental health professionals would be as influenced by hindsight bias, and, if they were, how this might have an impact on their ability to give objective expert witness testimony.

We developed a study to examine whether psychiatrists performing case reviews and estimating the risk of suicide or violence would provide responses suggestive of hindsight bias (i.e., exaggerate their ability to predict the adverse outcome). Our primary hypothesis was that psychiatrists performing case reviews who were provided with advance knowledge that an adverse outcome had occurred would make responses suggestive of hindsight bias (elevated responses on suicide or violence risk measures) when compared with their colleagues who reviewed cases without knowing the adverse outcome. We further hypothesized that nonforensic psychiatrists would be more prone to respond with hindsight bias than would forensic psychiatrists and that psychiatrists who were informed of an adverse outcome before case review would be more likely to rate care as having failed to meet the standard of care.

Methods

Participants

The investigators recruited subjects from a sample of psychiatrists living in the United States, who had an e-mail address listed in either the 2004 membership directory of the American Academy of Psychiatry and the Law (AAPL) or the online member directory of the American Psychiatric Association (APA). To create an evenly distributed geographic sample, we selected up to 14 members from each state and the District of Columbia who had listed e-mail addresses, by proceeding through the member lists alphabetically. When states did not have 14 members who met the inclusion criteria, the maximum number of members who met the criteria was selected. We cross-referenced the APA and AAPL directories to ensure that members recruited from the APA directory were not also AAPL members, so that APA members essentially represented APA, non-AAPL members (hereafter referenced as "APA-only" members for simplicity). Members listed as psychiatry residents or medical students in the APA list or who self-identified as psychiatry residents or medical students on the survey were excluded from data analysis.

Web-Based Informed Consent and Data Collection

We used e-mail invitations to recruit subjects. Within each e-mail message sent to potential participants, subjects were provided a hyperlink that directed them to a website with an electronic consent form for participation in the study. Consent was required before proceeding to the web-based survey. We informed potential participants that the purpose of the study was to learn more about psychiatrists' thinking regarding risk assessments and standard-of-care determinations. They were not told that hindsight bias was a factor under study. The study was approved by the University of Massachusetts Medical School Institutional Review Board.

We utilized a web-based survey instrument and followed published recommendations for data collection to ensure the validity and security of the data.[12–17] We used a commercially available software package and data collection service[18] to facilitate data collection in a valid and secure manner. Security features included the use of 128-bit, secure socket layer encryption to prevent interception of data during transmission from subjects' computers to the survey company's server and during downloading of data into the investigators' computers.

The use of unique identifiers further protected the confidentiality of potential participants. Identifiers were assigned to each potential participant and included in each survey invitation hyperlink. This method allowed us to keep each subject's identifiable personal data off the internet. By tracking each indi-

vidual's responses, we were able to minimize the risk that participants might submit more than one survey or forward their invitations to others, or that a random person browsing the internet might encounter and take the survey, each of which could potentially corrupt the data set.

Measures

Following provision of informed consent, we asked subjects to provide information on the following background characteristics: (1) years since completion of psychiatry residency, (2) membership status in APA and/or AAPL, (3) approximate percentage of time engaged in clinical psychiatry work, (4) approximate percentage of time engaged in forensic psychiatry work, (5) approximate population of the geographic area where subject practiced, and (6) gender of subject.

Subjects then reviewed two hypothetical cases of patients presenting for psychiatric care and the accompanying documentation, which was intentionally kept concise (under 750 words for each case) to encourage completion of the surveys. The first case involved a potentially suicidal patient, a 34-year-old man who presented to an emergency room with the chief complaint of having suicidal thoughts. A psychiatrist assessed the patient and released him for outpatient care with follow-up in 7 days, rather than admitting the patient to the hospital for inpatient treatment. The second case involved a potentially violent patient, a 20-year-old male psychiatric inpatient who initially presented to the hospital with homicidal ideation, was voluntarily admitted, and then signed a petition for release after 12 days of treatment. In this case, the psychiatrist decided to release the patient for outpatient care rather than file for commitment and continue inpatient hospitalization. To ensure that the hypothetical cases were representative of care that one might encounter in the community, we circulated the cases to 11 psychiatrists from three academic medical centers. They reviewed the cases for content and provided feedback, which was incorporated into the final case descriptions.

To test our hypotheses regarding hindsight bias, psychiatrist participants were randomized to a hindsight group and a control group. Participants in the hindsight group were provided with the outcome of care for each hypothetical case in advance of reviewing the cases and answering questions, while the control group participants were not. In the potentially suicidal patient case, we informed hindsight group participants that the patient committed suicide by shooting himself with a gun two days after his release from the emergency room. In the case of the potentially violent patient, we informed the hindsight group participants that the patient stabbed and killed a man five days after the patient was released from the hospital. Except for providing hindsight group participants with advance knowledge of case outcomes, the case materials provided to hindsight and control subjects were identical in every aspect.

Following participants' review of each case, they answered two closed-end questions. All subjects, regardless of group assignment, were reminded before answering the survey questions to formulate opinions based only on the information in the hospital chart. The first question was designed to garner psychiatrists' assessment of the risk of suicide or violence posed by the hypothetical patients at the time they were released from care on a scale of one to seven (with one denoting the lowest risk of suicide or violence, four a moderate risk, and seven the highest risk). Participants were then asked whether the assessment and care provided by the psychiatrist in each case met the standard of care and were given the chance to explain their reasoning.

Data Analysis

Comparisons of study groups (hindsight versus control) on background characteristics were made with chi-square tests for nominal measures, Mann-Whitney tests for ordinal measures, and analysis of variance (ANOVA) for continuous measures. Risk assessment ratings were treated as continuous data for the purposes of analysis. Multivariate analyses of risk assessment ratings were conducted by using generalized linear models (GLMs). Logistic regression was employed for multivariate analyses of standard-of-care determinations. Study group was always included in the models as were background variables associated ($p < .05$) with the risk assessment ratings and the corresponding interaction terms.

Results

Response Rate to Online Survey

We initially sent out e-mail invitations to 1,371 psychiatrists. Of those, 401 (29.2%) were returned as undeliverable. This left 970 presumably deliverable invitations to active e-mail accounts. For potential

Table 1 Background Characteristics of Study Sample

Background Characteristics	Study Group: Hindsight (n = 114)	Study Group: Control (n = 121)	Total (n = 235)	p*
Male (%)	70.2	72.7	71.5	.665
Years since completion of residency (%)				
0–10	40.4	27.3	33.6	.054
11–20	21.1	27.3	24.3	
21–30	24.6	24.0	24.3	
>30	14.0	21.5	17.9	
Membership status, %				
APA only	53.5	47.9	50.6	.393
AAPL	46.5	52.1	49.4	
Median percentage of time spent on clinical work (IQR†)	70 (50)	80 (50)	80 (50)	.411
Median percentage of time spent on forensic psychiatry work (IQR†)	10 (30)	10 (45)	10 (40)	.111
Practices in a geographic area with a population of ≥1 million (%)	41.2	46.3	43.8	.435

*Chi-square test used for nominal measures and Mann-Whitney test for ordinal measures.
†Interquartile range is the difference between the 75th and the 25th percentiles and is a robust estimator of a distribution's dispersion or spread.

participants who did not respond to the first e-mail invitation, we sent a second, follow-up e-mail invitation within two weeks of the initial invitation. Two hundred and thirty-six subjects responded to the survey, yielding a 24.3 percent response rate, comparable with the response rate of similarly designed web-based surveys reported in the literature.[12–14] One participant was excluded from data analysis due to self-identifying as a psychiatry resident.

Sample Description

The hindsight and control groups were balanced in size (n = 114 and 121, respectively). The majority (71.5%) of the sample was male (Table 1). Less than half (43.8%) reported practicing in geographic areas of more than 1 million persons. The median time since completing residency was 11 to 20 years. The median amount of time devoted to clinical work was 80 percent. The median amount of time engaged in forensic work was 10 percent. About half (50.6%) of the study participants were APA-only members and the remainder were AAPL members. Comparisons between the hindsight and control groups on these measures revealed no statistically significant differences. Regarding time since residency, there was a trend for a higher proportion of respondents in the hindsight group to report having completed their residencies within the past 10 years relative to the control group (p = .054).

Comparison of the APA-only group and the AAPL group showed no significant differences in gender, years since completion of residency, or the proportion of individuals practicing in a geographic area of more than 1 million persons. There were significant differences in the median times spent performing clinical and forensic work among these groups, with APA-only members spending a significantly greater amount of time engaged in clinical work, and AAPL members devoting a significantly greater amount of time to forensic work (Table 2).

Participants' Estimate of Suicide and Violence Risk

Consistent with the primary hypothesis of the study, hindsight subjects as a group endorsed significantly higher suicide and violence risk ratings than did control subjects (Table 3). With respect to examining for hindsight bias based on background characteristics, we observed effects based on organization membership status. Both APA-only members and AAPL members endorsed higher suicide risk ratings

Table 2 Background Characteristics by Organization Affiliation

Background Characteristics	Study Group: APA-only (n = 119)	Study Group: AAPL (n = 116)	Total (n = 235)	p
Male (%)	68.1	75.0	71.5	.239
Years since completion of residency (%)				
0–10	36.1	31.0	33.6	.665
11–20	21.0	27.6	24.3	
21–30	26.1	22.4	24.3	
>30	16.8	19.0	17.9	
Median percentage of time spent on clinical work (IQR)	90 (50)	70 (50)	80 (50)	.000
Median percentage of time spent on forensic psychiatry work (IQR)	0 (10)	20 (40)	10 (40)	.000
Practices in a geographic area with a population of ≥1 million (%)	38.7	49.1	43.8	.105

Data are as described in Table 1.

Table 3. Suicide and Violence Risk and Standard of Care Assessment Results by Study Group

	Study Group			
	Hindsight	Control	Total	*p**
Risk assessment rating†				
Mean suicide risk rating (SD)	(*n* = 114)	(*n* = 117)	(*n* = 231)	.000
	4.9 (1.04)	4.2 (1.16)	4.6 (1.2)	
Mean violence risk rating (SD)	(*n* = 112)	(*n* = 115)	(*n* = 227)	.008
	4.0 (1.12)	3.6 (1.16)	3.8 (1.2))	
Met standard of care‡				
Suicide (%)	(*n* = 114)	(*n* = 117)	(*n* = 231)	.113
	49.1	59.0	54.1	
Violence (%)	(*n* = 112)	(*n* = 115)	(*n* = 227)	.251
	67.0	73.9	70.5	

*ANOVA was used for continuous measures and the chi-square test for nominal measures.
†Higher rating represents higher perceived risk on a seven-point Likert scale.
‡Responded "yes."

in hindsight than did their control counterparts, but the difference between hindsight and control groups was statistically significant for APA-only members whereas it was not for AAPL members (Table 4). We also observed a significant difference between the APA-only and AAPL group in the control condition, with the AAPL group showing a significantly higher mean suicide risk rating than did the APA-only group ($p = .015$). Multivariate analysis revealed that this interaction effect between study group and organization membership status was statistically significant ($p = .010$).

In the violence case, we observed a similar hindsight bias effect based on membership status. The APA-only members endorsed significantly higher violence risk ratings in hindsight than did their control counterparts, but AAPL members did not (Table 4). In the violence case, we also observed a difference between the APA-only and AAPL group in the control condition, with the AAPL group providing a significantly higher mean violence risk rating than did the APA-only group ($p = .041$). Multivariate analysis showed significant main effects of study group ($p = .006$) and AAPL membership status ($p = .030$).

Table 4 Suicide and Violence Risk and Standard of Care Assessment Results by Study Group and AAPL Status

	Study Group		
	Hindsight	Control	*p**
Risk assessment rating†			
Mean suicide risk rating (SD)			
APA only	(*n* = 61)	(*n* = 57)	.000
	5.0 (1.1)	4.0 (1.3)	
AAPL	(*n* = 53)	(*n* = 60)	.118
	4.8 (1.0)	4.5 (1.0)	
Mean violence risk rating (SD)			
APA-only	(*n* = 60)	(*n* = 57)	.011
	3.9 (1.2)	3.4 (1.1)	
AAPL	(*n* = 52)	(*n* = 58)	.173
	4.1 (1.1)	3.8 (1.2)	
Met standard of care‡			
Suicide (%)			
APA-only	(*n* = 61)	(*n* = 57)	.092
	45.9	61.4	
AAPL	(*n* = 53)	(*n* = 60)	.683
	52.8	56.7	
Violence (%)			
APA-only	(*n* = 60)	(*n* = 57)	.068
	70.0	84.2	
AAPL	(*n* = 52)	(*n* = 58)	.971
	63.5	63.8	

Data are as described in Table 3.

Participants' Opinions Regarding the Standard of Care

For both suicide and violence conditions, subjects on the whole did not provide responses suggestive of hindsight bias on measures of standard of care (Table 3). In the suicide case, we observed a significant bivariate association between gender and standard-of-care rating, with women being more likely than men to rate care as negligent ($p = .035$). Multivariate analysis revealed no other significant main effects and no significant interaction effects. In the suicide case, we observed a trend for the APA-only group to rate care as negligent in hindsight more frequently, but the difference was not statistically significant ($p = .092$).

In the violence case, APA-only members displayed a tendency to rate care as negligent in hindsight more frequently (Table 4), although the difference was of borderline statistical significance ($p = .068$). In addition, we observed a significant difference between the APA-only and AAPL group in the control condition, with the AAPL group significantly more likely to rate care as negligent than was the APA-only group ($p = .013$); this difference was not present in hindsight. Multivariate analysis showed that the main effect of organization membership status was statistically significant ($p = .015$).

Discussion

The current study supported the investigators' primary hypothesis that psychiatrists performing case reviews who were provided with advance knowledge of an adverse outcome would offer responses suggestive of hindsight bias. This finding is consistent with previous studies on hindsight bias, one among anesthesiologists reviewing cases with adverse anesthetic outcomes[11] and one among mock jurors reviewing therapist actions in *Tarasoff*-type cases in which violence had occurred.[5]

Caution is indicated in generalizing the results of the current study based on the use of brief, hypothetical case scenarios and the use of a web-based survey format. However, the study highlights the importance of the potentially biasing effect of knowledge of an adverse outcome on expert opinion. This finding has potential relevance to those providing expert opinions in malpractice cases involving suicide or violence. If experts unknowingly exaggerate the risk of suicide or violence following a patient encounter, they could overestimate the causal role of the actions or omissions of the treating psychiatrist.

Although hindsight bias on measures of risk assessments was strong in the current study, the findings with regard to dichotomous determinations of negligence were not significant—a somewhat reassuring result that suggests that ultimate determinations of negligence in malpractice cases may be less affected by hindsight bias than estimates of violence or suicide risk. However, in the APA-only study group, we observed trends in both the suicide and violence scenarios for the hindsight group to rate care as negligent more frequently (Table 4). Retrospective power analysis revealed that almost three times as many APA-only respondents would have been necessary in both study groups for the observed differences to reach statistical significance, which limits the certainty with which we can conclude that the APA-only group did not truly engage in responses suggestive of hindsight bias.

Perhaps the most interesting finding was the lack of suggestion of hindsight bias among the AAPL member group. The findings suggest that those belonging to a professional organization that promotes ongoing education and research in the area of psychiatry and the law and the ethical practice of forensic psychiatry were buffered against the effects of hindsight bias. A significantly greater percentage of AAPL psychiatrists compared with APA-only psychiatrists reported engaging in expert work (forensic practice), and this may be a reason for a lack of responses suggestive of hindsight bias in the AAPL group as a meta-analysis on hindsight bias studies found smaller hindsight bias effect sizes for studies in which participants were experts or familiar with the task under evaluation compared with nonexperts.[19] Although a more recent meta-analysis did not replicate this finding,[20] significantly greater involvement in forensic practice or forensic training may promote a more standardized approach to suicide and violence risk assessments that leads to opinions derived from case content, rather than outcomes. It could also be that the very nature of forensic work, in which forensic psychiatrists' opinions are routinely scrutinized by attorneys, judges and juries, and other experts, may lead to greater pause and emphasis on case content, combating hindsight bias.

To our knowledge, this study is the first of its kind to use a randomized, controlled format to investigate the phenomenon of bias among psychiatrists providing opinions as they might during expert work. The format could be used to elucidate whether psychiatrists are susceptible to other forms of bias in forensic work, such as bias due to countertransference, dual-agency affiliation, attorney referral source, or financial incentive. If further studies demonstrate the potential for bias in expert work from such sources, then debiasing strategies should be sought. The courts have developed such strategies in attempts to debias jurors, including techniques used in attorneys' closing arguments,[7] special juror instructions, and bifurcated hearings.[21] Similarly, specific training for expert witnesses in threats to the objectivity of medicolegal evaluations may be helpful in this context, but remains to be tested empirically.

References

1. American Academy of Psychiatry and the Law: Ethics Guidelines for the Practice of Forensic Psychiatry. Adopted May 2005. Available at http://www.aapl.org/pdf/ETHICSGDLNS.pdf. Accessed January 23, 2006
2. Commons ML, Miller PM, Gutheil TG: Expert witness perceptions of bias in experts. J Am Acad Psychiatry Law 32:70–5, 2004
3. Sattar SP, Pinals DA, Gutheil TG: Countering countertransference: a forensic trainee's dilemma. J Am Acad Psychiatry Law 30:65–9, 2002
4. Goldstein AS: The psychiatrist and the legal process: the proposals for an impartial expert and for preventive detention. Am J Orthopsychiatry 33:123–31, 1963
5. LaBine SJ, LaBine G: Determinations of negligence and the hindsight bias. Law Hum Behav 20:501–16, 1996
6. Baron J, Hershey JC: Outcome bias in decision evaluation. J Personality Soc Psychol 54:569–79, 1988
7. Stallard JS, Worthington DL: Reducing the hindsight bias utilizing attorney closing arguments. Law Hum Behav 22:671–83, 1998
8. Hugh TB, Tracy GD: Hindsight bias in medicolegal expert reports. Med Law 176:277–8, 2002
9. Arkes HR, Wortmann RL, Saville PD, *et al*: Hindsight bias among physicians weighing the likelihood of diagnoses. J Appl Psychol 66:252–4, 1981
10. Sacchi S, Cherubini P: The effect of outcome information on doctors' evaluations of their own diagnostic decisions. Med Educ 38:1025–7, 2004
11. Caplan RA, Posner KL, Cheney FW: Effect of outcome on physician judgments of appropriateness of care. JAMA 265:1957–60, 1991
12. Sills SJ, Song C: Innovations in survey research: an application of web-based surveys. Soc Sci Computer Rev 20:22–9, 2002
13. Smee A, Brennan M: Electronic surveys: a comparison of e-mail, web and mail. ANZMAC 2000 Visionary Marketing for the 21st Century: Facing the Challenge. Available at http://130.195.95.71:8081/www/anzmac2000/cdsite/papers/s/smee1.pdf. Accessed January 1, 2005
14. Brawner CE, Felder RM, Allen RH, *et al*: A Comparison of electronic surveying by e-mail and web. 2001 ASEE Annual Conference Proceedings, June 2001. Available at http://www.ncsu.edu/felder-public/papers/web-email_surveys.pdf. Accessed January 1, 2005
15. Kraut RE, Olson J, Manaji M, *et al*: Psychological research online: report of Board of Scientific Affairs' Advisory Group on the conduct of research on the internet. Am Psychol 59:105–17, 2004
16. Schaefer DR, Dillman DA: Development of a standard e-mail methodology: results of an experiment. Public Opin Q 62:378–97, 1998
17. Best SJ, Krueger B, Hubbard C, *et al*: An assessment of the generalizability of internet surveys. Soc Sci Computer Rev 19:131–45, 2001
18. SurveyMonkey. Available at http://www.surveymonkey.com. Accessed March 3, 2005
19. Christensen-Szalanski JJJ, Willham CF: The hindsight bias: a meta-analysis. Organ Behav Hum Decis Process 48:147–68, 1991
20. Guilbault RL, Bryant FB: A meta-analysis of research on hindsight bias. Basic Appl Soc Psychol 26:103–17, 2004
21. Smith AC, Greene E: Conduct and its consequences: attempts at debiasing jury judgments. Law Hum Behav 29:505–26, 2005