

2001

Statistics, Not Experts

Cass R. Sunstein

William Meadow

Follow this and additional works at: http://chicagounbound.uchicago.edu/journal_articles



Part of the [Law Commons](#)

Recommended Citation

Cass R. Sunstein & William Meadow, "Statistics, Not Experts," 51 Duke Law Journal 629 (2001).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

Essay

STATISTICS, NOT EXPERTS

WILLIAM MEADOW†

CASS R. SUNSTEIN††

INTRODUCTION

Legal proceedings, both judicial and administrative, frequently raise questions about what people, including experts, ordinarily do.¹ If

Copyright © 2001 by William Meadow and Cass R. Sunstein.

† Associate Professor, Pediatrics; Associate Director, Neonatology; Assistant Director, MacLean Center for Clinical Medical Ethics, University of Chicago.

†† Karl N. Llewellyn Distinguished Service Professor of Jurisprudence, Law School and Department of Political Science, University of Chicago.

We are grateful to Gertrud Fremling, James Heaton, Anup Malani, John Monahan, and Richard A. Posner for valuable comments on an earlier draft. Laura Warren provided excellent research assistance and valuable comments.

1. There is a dispute about the weight to be given to ordinary practice. Generally in negligence actions, custom is considered informative but not definitive. RICHARD A. EPSTEIN, *CASES AND MATERIALS ON TORTS* 201–42 (7th ed. 2000); PROSSER AND KEETON ON THE LAW OF TORTS 193, 195 (W. Page Keeton ed., 5th ed. 1984) [hereinafter PROSSER AND KEETON]:

[E]vidence of the usual and customary conduct of others under similar circumstances is normally relevant and admissible, as an indication of what the community regards as proper. . . . [A]s a general rule, the fact that a thing is done in an unusual manner is merely evidence to be considered in determining negligence, and is not in itself conclusive.

For the classic case, see *The T.J. Hooper*, 60 F.2d 737, 740 (2d Cir. 1932) (holding that conformity to custom does not preclude a finding of negligence). See also *Trimareo v. Klein*, 436 N.E.2d 502, 506 (N.Y. 1982) (refusing to give custom in shower door installation conclusive weight).

In many jurisdictions, however, medical malpractice is an exception to the rule, and custom defines reasonable conduct. EPSTEIN, *supra*, at 225 (identifying the role of custom in medical malpractice cases as “conclusive on the standard of medical care”); PROSSER AND KEETON, *supra*, at 189 (“[T]he standard of conduct becomes one of ‘good medical practice,’ which is to say, what is customary and usual in the profession.”); see, e.g., *Burgess v. Superior Court*, 831 P.2d 1197, 1206 (Cal. 1992) (holding that the legal standard of care required for doctors is the standard of practice required by their own profession); *Cross v. Huttenlocher*, 440 A.2d 952, 954 (Conn. 1981) (“A physician is under a duty to his patient to exercise that degree of care, skill and diligence which physicians in the same general line of practice ordinarily pos-

a doctor is accused of negligence, for example, it is necessary to know about the customary practice of doctors.² Of course, negligence judgments depend, at least in part, on an assessment of ordinary practice.³ But how is ordinary practice by professionals or others to be assessed?

In trials and on appeal, the basic answer is that the assessment comes via statements from expert witnesses who describe the ordinary practice.⁴ There can be no doubt that experts know a great deal about topics on which ordinary people lack information. But experts, no less than other people, are subject to predictable biases.⁵ Their judgments about risk are affected by the same heuristics and biases to

sess and exercise in similar cases.”); *Spensieri v. Lasky*, 723 N.E.2d 544, 548 (N.Y. 1999) (stating that a physician will usually be insulated from tort liability where there is evidence that he or she conformed to accepted community standards of practice); *Toth v. Cmty. Hosp.*, 239 N.E.2d 368, 372 (N.Y. 1968) (“The law generally permits the medical profession to establish its own standard of care.”); *MacPherson v. Ellis*, 287 S.E.2d 892, 896 (N.C. 1982) (holding that failure to adhere to customary practices of informed consent results in civil liability); *Boone v. Fisher*, No. 13-96-001-CV, 2000 Tex. App. LEXIS 4244, at *4 (Tex. App. June 22, 2000) (stating that a trier of fact, when determining whether a doctor’s actions were reasonable, must be advised of standard medical practice). *But cf.* *McNeill v. United States*, 519 F. Supp. 283, 288 (D.S.C. 1981) (cautioning that customary medical treatment is not conclusive on the exercise of care); *Vassos v. Roussalis*, 625 P.2d 768, 772 (Wyo. 1981) (“[T]he skill, diligence, knowledge, means and methods are not those ‘ordinarily’ or ‘generally’ or ‘customarily’ exercised or applied, but are those that are ‘reasonably’ exercised or applied.”).

2. See *supra* note 1.

3. We build on existing law here, without intending to enter into debates about the extent to which law should simply incorporate, or improve upon, the existing standard of care. See *supra* note 1; see also, e.g., *Advincula v. United Blood Servs.*, 678 N.E.2d 1009, 1027 (Ill. 1996) (suggesting that existing custom is not determinative); *Roach v. Springfield Clinic*, No. 73,394, 1992 Ill. LEXIS 204, at *33–*36 (Ill. Dec. 4, 1992) (same); *Darling v. Charleston Cmty. Mem’l Hosp.*, 211 N.E.2d 253, 257 (Ill. 1965) (same). On any view, the customary practice is relevant, and that is sufficient for our purposes here.

4. See *supra* note 1.

5. See Hillel Einhorn, *Expert Judgment: Some Necessary Conditions and an Example*, in JUDGMENT AND DECISION MAKING 324, 325 (Terry Connolly et al. eds., 2d ed. 2000) (discussing the decisional factors and biases of expert judgment); David Faust & Jay Ziskin, *The Expert Witness in Psychology and Psychiatry*, in JUDGMENT AND DECISION MAKING, *supra*, at 336, 341–44 (discussing factors that limit clinical judgments); Gary Gaeth & James Shanteau, *Reducing the Influence of Irrelevant Information on Experienced Decision Makers*, in JUDGMENT AND DECISION MAKING, *supra*, at 305, 306–16 (outlining empirical evidence of experts’ use of irrelevant information in decisionmaking); Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 3, 18 (Daniel Kahneman et al. eds., 1982) (presenting the intuitive bias of statistical experts).

which most people are subject, even if (and this is a disputed question) expertise tends to reduce the most serious errors.⁶

To correct the resulting problems, we offer a simple proposal: The legal system should rely, whenever it can and far more than it now does, on *statistical data* about doctors' performance rather than on the opinions of experts about doctors' performance. For the first time, it is becoming possible for law to rely on this evidence, for the simple reason that it is becoming increasingly available. If our argument is convincing in the medical context, it should apply in many other settings in which experts are asked to testify about negligence or deviations from ordinary practices. In many settings, the fallible opinions of isolated experts should be supplemented or replaced by statistical data. Those opinions should be seen as a kind of crude second-best, far inferior to the data that it approximates. A step toward reliance on such data would dramatically increase the sense and rationality of tort law, and perhaps other areas of law as well.

We are hopeful that the use of statistical data will simplify some of the central issues in litigation and reduce the role of strategic behavior in the litigation process. Battles between isolated experts will be easier to mediate when there is a large pool of evidence on which to draw. But our larger claim is that by using statistical data, the legal system will reach far more accurate results. If the law is seeking to determine the standard of care, it should not depend on fallible memories and recollections of local experiences. It should draw instead on more global evidence of the kind that is increasingly available for use in court and elsewhere.⁷

6. See Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 781-811 (2001) (discussing errors in risk-related judgments by judges).

7. It should be clear that this Essay is intended as a contribution to the rapidly growing work in what is sometimes called behavioral law and economics. See, e.g., Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998) (proposing that law and economics analysis can be improved by increased attention to patterns of actual human behavior). Our proposal also is in line with recent work on the general value of using social science in law. For a valuable discussion of the increased use of social science evidence in law, see generally Laurens Walker & John Monahan, *Social Frameworks: A New Use of Social Science in Law*, 73 VA. L. REV. 559 (1987).

I. EXCESSIVE OPTIMISM ABOUT RISKS

A. *In General*

It is now well known that most normal people tend to be “risk optimists,” in the particular sense that they believe themselves to be relatively insulated from risks that are faced by others who are similarly situated.⁸ This is one of the most robust findings in cognitive psychology. For example, ninety percent of drivers believe that they are safer than most drivers and less likely to be involved in a serious accident.⁹ Most people believe that they are distinctly unlikely to be subject to various risks, such as cancer, heart disease, and divorce.¹⁰ In some important domains, people appear not only to believe themselves less likely than others to fall prey to certain risks, but also to underestimate the actual risk to which they are subject.¹¹ With respect to automobile accidents, people believe that the danger they face is less than it is as a matter of statistical fact.¹² Smokers appear to know

8. Christine Jolls, *Behavioral Economic Analysis of Redistributive Legal Rules*, in *BEHAVIORAL LAW AND ECONOMICS* 288, 290–92 (Cass R. Sunstein ed., 2000) (discussing excessive optimism); Neil D. Weinstein, *Unrealistic Optimism About Future Life Events*, 39 *J. PERSONALITY & SOC. PSYCHOL.* 806, 807–12 (1980) (discussing individuals’ tendency to believe themselves to be less vulnerable than other people to adverse events).

W. Kip Viscusi, in his article *Using Warnings to Extend the Boundaries of Consumer Sovereignty*, 23 *HARV. J.L. & PUB. POL’Y* 211 (2000), offers an intriguing qualification, suggesting that although people might believe that they are less subject to certain risks than most other people, they also might have a statistically accurate understanding of the risks to which they are subject. *Id.* at 227. In other words, they might believe that they are less likely to have a car accident than most people, but also have an accurate sense of the lifetime risk that they will die in a car crash (say, 1 in 1000). *Id.* Viscusi offers evidence to support this hypothesis. *Id.* at 216–17 (describing evidence of accurate risk assessments of workers). *But see* John Ayanian & Paul Cleary, *Perceived Risks of Heart Disease and Cancer Among Cigarette Smokers*, 281 *JAMA* 1019, 1020–21 (1999) (suggesting an underestimate of the actual risk for smokers); Jolls, *supra*, at 291 (same in the context of car accidents). In any event, resolving this debate is not necessary for our purposes. The very possibility of expert error and overoptimism, confirmed by the studies described in this Essay, is enough to justify reliance on statistics rather than on isolated experts.

9. SHELLEY E. TAYLOR, *POSITIVE ILLUSIONS* 10–11 (1989).

10. Weinstein, *supra* note 8, at 810.

11. See Jolls, *supra* note 8, at 291 (“[P]eople [underestimate] their own probability as opposed to overestimating the average person’s probability.”).

12. See Richard Arnould & Henry Grabowski, *Auto Safety Regulation: An Analysis of Market Failure*, 12 *BELL J. ECON.* 27, 34–35 (1981) (discussing excessive optimism about the risk of automobile accidents); Colin Camerer & Howard Kunreuther, *Decision Processes for Low Probability Events: Policy Implications*, 8 *J. POL’Y ANALYSIS & MGMT.* 565, 566–67 (1989) (same).

the statistical risks of smoking,¹³ but they believe that they are less likely than most smokers to fall victim to the various risks.¹⁴ In one study, less than half of smokers surveyed believed that they had a higher-than-average risk of cancer or cardiovascular disease; indeed, most heavy smokers (at least forty cigarettes per day) believed that they were not at any increased risk.¹⁵ Only one group of people does not show a tendency to excessive optimism: the clinically depressed.¹⁶

It would be natural to infer that unrealistic optimism, in many cases, is adaptive. Though such optimism may cause false predictions of outcomes, an optimistic attitude may increase the probability of a good outcome and otherwise create hedonic benefits.¹⁷ This may be particularly true in the medical context; perhaps an excessively optimistic doctor is more likely to get good results. If so, optimism is well adapted to the social role of a doctor. It also might be thought that some social settings work against optimistic bias. In markets, for example, entrepreneurs might have an incentive toward realism, especially because market pressures—it might be expected—would punish and drive out those who are unrealistically optimistic. But even if markets move people toward realism, entrepreneurs, no less than anyone else, have been shown to suffer from unrealistic optimism.¹⁸ It is noteworthy that expert witnesses are not disciplined by ordinary market pressures.¹⁹ If experts err, they might even help their clients,

13. See W. KIP VISCUSI, *SMOKING* 4 (1992) (showing substantial awareness among individuals of the statistical risks of smoking).

14. Ayanian & Cleary, *supra* note 8, at 1020-21.

15. *Id.*

16. TAYLOR, *supra* note 9, at 212-15, 226 (noting the lack of excessive optimism among depressed people); Peter Lewinsohn et al., *Social Competence and Depression*, 89 J. ABNORMAL PSYCHOL. 203, 207-08, 210-11 (1980) (same).

17. This is the thesis of TAYLOR, *supra* note 9. For a discussion of optimism and self-fulfilling prophecies, see NICHOLAS CHRISTAKIS, *DEATH FORETOLD: PROPHECY AND PROGNOSIS IN MEDICAL CARE* 135-62 (1999) (evaluating the role of self-fulfilling prophecies in medicine). There are analogies in the domain of athletics. Consider the comments of basketball coach Lenny Wilkins: "The shooter should think the shot's going in. Everyone else should be thinking he's going to miss." Phil Jasner, *Sixers Find the Road Better Than Home*, PHILA. DAILY NEWS, Dec. 30, 2000, LEXIS, Phila. Daily News File.

18. See KENNETH R. MACCRIMMON & DONALD A. WEHRUNG, *TAKING RISKS* 7, 260 (1986) (discussing the optimism of entrepreneurs); Daniel Kahneman & Dan Lovallo, *Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking*, 39 MGMT. SCI. 17, 27-28 (1993) (same).

19. In this respect, experts are like juries, which also lack ordinary market discipline. Of course there is something of a market for expert witnesses, so that people will lack credibility, and opportunities to be witnesses, if they are systematically (shown to be) wrong. But it would be heroic to suggest that these pressures ensure accuracy from experts.

and hence excessive optimism might well be favored, rather than disfavored, by whatever market pressures exist in this domain.

B. *Optimism and Pessimism in Clinical Practice*

It seems reasonable to speculate that those with specialized knowledge are far less prone to optimistic bias and to other motivational and cognitive errors. Unfortunately, this appears not to be the case. It has been found almost universally that physicians, no less than other people, are prone to make cognitive mistakes and indeed have a substantial tendency to err in predicting outcomes.²⁰ "Virtually all" of the existing studies of physicians "have documented frequent and large errors in predictions."²¹ No study finds a high level of accuracy. The errors tend in a particular direction: "physicians are prone to an optimistic bias."²²

As an example, consider a study published in 1972 that showed that in making predictions about the length of survival for cancer patients, only 47% of the physicians provided prognoses that were even roughly accurate and that 80–90% of the mistakes were excessively optimistic.²³ A subsequent study published in 1987 found overestimates of survival time in 88% of the cases, by roughly a factor of three.²⁴ A recent, large-scale study found inaccurate predictions in 80% of cases, with 63% of these showing overestimates.²⁵ Interestingly, patient characteristics (age, race, sex, illness duration) were not correlated with inaccurate, optimistic estimates, but one factor was:

20. For a general discussion of this finding, see CHRISTAKIS, *supra* note 17, at 66–71 (describing studies that conclude that physicians tend to substantially overestimate survival); on doctors and patients, see Donald Redelmeier et al., *Understanding Patients' Decisions*, 270 JAMA 72, 72–76 (1993) (encouraging physicians to recognize that patients often make irrational medical decisions and that it is physicians' role to identify pitfalls in reasoning).

21. CHRISTAKIS, *supra* note 17, at 66.

22. *Id.*

23. *Id.* (describing C. Murray Parkes, *Accuracy of Predictions of Survival in Later Stages of Cancer*, BRIT. MED. J., Vol. 2, 1972, at 29, 29–31 which examined London physicians' prognoses for 168 cancer patients, where the prognostic accuracy was defined as no more than twice and no less than half the actual duration of survival).

24. *Id.* at 66–67 (describing Louis H. Heyse-Moore & V.E. Jolmson-Bell, *Can Doctors Accurately Predict the Life Expectancy of Patients with Terminal Cancer?*, 1 PALLIATIVE MED. 165, 165–66 (1987), which found that physicians' predicted survival rates for terminal cancer patients averaged eight weeks, compared to an observed average survival rate of two weeks).

25. *Id.* at 67, 201–08 (discussing Nicholas A. Christakis & Elizabeth B. Lamont, *The Extent and Determinants of Error in Physicians' Prognoses for Terminally Ill Patients* (unpublished manuscript) and comparing the mortality predictions of physicians for 504 Chicago patients admitted during 1996 to observed outcomes).

“the better the doctor knew the patient—as measured, for example, by the length and intensity of their contact—the more likely the doctor was to err in the prognosis, most frequently by overestimating survival.”²⁶ The overall picture is that in cases involving cancer patients, physicians accurately predict survival only 10–30% of the time, and the rest of the time they overestimate survival by a factor of two to five.²⁷

All of this is highly suggestive but not decisive with respect to the particular point that we mean to urge here: that experts make erroneous judgments about the ordinary standard of medical care, and that they tend to err in a particular direction. Because expert witnesses in medical malpractice cases typically are doctors, it would be plausible to think that their general tendency toward optimism with their patients would affect their testimony as experts, causing them to overstate systematically the efficiency and effectiveness of ordinary practice. At the same time, it also would be possible to imagine that (a) doctors’ predictions about their patients’ prospects are systematically overoptimistic, but (b) doctors have an accurate sense of what is ordinarily done by themselves and other doctors. If there is no systematic error with respect to (b), there would be little need to substitute statistical data for the expert testimony of doctors. We now offer the first real evidence in the legal literature on the central question, showing that medical experts systematically err.²⁸

26. *Id.* at 67.

27. *Id.* at 68.

28. In a separate phenomenon, physicians are sometimes informally accused of deliberately offering excessively optimistic or pessimistic prognoses in their clinical practices. Both of these behaviors are understandable in context. Excessively pessimistic prognoses, commonly referred to as “hanging crepe,” usually accompany discussions of procedures with particularly problematic outcomes (e.g., surgery for aortic aneurisms, carotid endarterectomies, cardiopulmonary resuscitation of twenty-three-week-gestation infants). Mark Siegler, *Pascal’s Wager and the Hanging of Crepe*, 293 *NEW ENG. J. MED.* 853, 853–57 (1975). In these cases, the underlying idea is that the patients should be fully prepared for the statistically expected result (a poor outcome or even death), and that in the unlikely, but possible, event of a good outcome, both physician and patient alike will be pleasantly surprised. *Id.* at 853–54. If patients are prone to optimistic bias, excessively pessimistic prognoses might even be justified as a possible corrective. There also are well-recognized examples of what appears to be excessive clinical optimism. End-of-life prognostications are notoriously “optimistic” in that the time to death is commonly overestimated. Nicholas A. Christakis & Elizabeth B. Lamont, *Extent and Determinants of Error in Doctors’ Prognoses in Terminally Ill Patients: Prospective Cohort Study*, 320 *BRIT. MED. J.* 469, 472 (2000). This error also is understandable in context—to the extent that there is any element of “self-fulfilling prophecy” in these situations, physicians may wish their patients to benefit from this effect.

II. STANDARD OF CARE: EXPERTS VERSUS DATA

A. *Bacterial Meningitis*

Consider the problem of bacterial meningitis in children.²⁹ For the purposes of the current discussion, it is necessary to know only three things about this disease and its treatment. Bacterial meningitis is an infection of the brain; it (usually) can be treated with antibiotics; and, as a general rule, the sooner the antibiotics are administered, the better the outcome will be.³⁰

In medical malpractice cases, the question often arises whether treatment of a particular child was unduly "delayed,"³¹ because delay can cause serious harm.³² Of course a key issue is how a legal decisionmaker would know what kind of delay counts as "undue." For

Both of these conditions are simply tangential to the claims made here. We are suggesting that expert descriptions of actual practice, not the discussions that accompany these practices, are often inaccurate, and predictably so. Practitioners may or may not be well motivated in their attempts to exaggerate outcomes in order to convince their patients of the desirability of a particular course of action *ex ante*. This may or may not be a good thing to do. But it is irrelevant to expert testimony. Our purpose here is to minimize the possibility that, when charged with describing the actual spectrum of clinical care, experts exaggerate their descriptions *ex post*.

29. Some of the data here is discussed for a nonlegal audience and from a very different angle in William Meadow et al., *Ought "Standard Care" Be the "Standard of Care"? A Study of the Time to Administration of Antibiotics in Children with Meningitis*, 147 AM. J. DISEASES CHILD. 40, 40-44 (1993) (concluding that "standard of care" is highly subjective and should be defined by data analysis rather than expert opinion).

30. Acute bacterial meningitis is one of the most life-threatening diseases of childhood and constitutes a medical emergency. 2 EMERGENCY MEDICINE: CONCEPTS AND CLINICAL PRACTICE 1213 (Peter Rosen ed., 4th ed. 1998). Death may occur within minutes to hours unless rapid diagnoses are made and appropriate antibiotics are administered immediately. ROSCOE N. GRAY & LOUISE J. GORDY, ATTORNEYS' TEXTBOOK OF MEDICINE § 33.62(1a) (Lois L. Caswell et al. eds., 3d ed. 1992). "Some of the immediate complications of bacterial meningitis include coma with loss of protective airway reflexes, seizures, cerebral edema, vasomotor collapse, disseminated intravascular coagulation, respiratory arrest, dehydration, pericardial effusion, and death." 3 EMERGENCY MEDICINE, *supra*, at 2202.

31. See, e.g., *N. Trust Co. v. Louis A. Weiss Mem'l Hosp.*, 493 N.E.2d 6, 12 (Ill. App. 1986) (holding that "evidence which shows . . . that negligent delay in diagnosis or treatment . . . lessened the effectiveness of treatment is sufficient to establish proximate cause" in medical malpractice cases).

32. There is a presumed correlation between the duration of symptoms before administration of antibiotics and the likelihood of subsequent permanent brain damage. The appropriate time scale (not minutes, perhaps hours, certainly days) to evaluate these time intervals is best addressed in Michael Radetsky, *Duration of Symptoms and Outcome in Bacterial Meningitis*, 11 PEDIATRIC INFECTIOUS DISEASE J. 694, 694-98 (1992). See also Meadow et al., *supra* note 29, at 43 (discussing how this presumed correlation breaks down when it is considered along the time scale of minutes).

obvious reasons, juries are assumed to be ignorant of the general medical facts in such cases, and consequently are instructed to rely on the testimony of expert witnesses. The instruction that jurors should rely on expert opinions for their knowledge about the standard of medical care answers one question, but it raises another—on what should the experts themselves rely?

In the traditional formulation, experts are instructed to rely on their personal “knowledge and training.” For the reasons stated, however, such admonitions raise many problems. In the context of bacterial meningitis, for example, it would be reasonable to predict that experts would tend to be excessively optimistic—that their recollections of the amount of time that elapsed before antibiotics were initiated in children with meningitis would be tilted toward shorter durations than actually occurred. And if this is the case, these experts, relying on their “knowledge and training,” will provide inaccurate accounts of behavior.

We attempted to test this hypothesis by surveying doctors about their experiences in cases of childhood meningitis and comparing these survey responses to detailed descriptions of physicians’ behaviors in actual cases of meningitis.³³ We began by identifying doctors with expertise in the two most relevant subspecialties of pediatrics, pediatric emergency room medicine (ER) and pediatric infectious diseases (ID). These are the fields from which potential experts most likely would be drawn to testify in lawsuits about deviations from the standard of care in meningitis cases. We presented these doctors with a questionnaire about a child who was brought to the emergency room with symptoms suggestive of bacterial meningitis. We asked how long, in their opinion, it would take to administer antibiotics in

33. Physicians were surveyed using an anonymous, voluntary questionnaire distributed at the national meeting of the Society for Pediatric Research in 1990. The respondents were identified only by specialty: general pediatrics, pediatric infectious diseases, pediatric emergency room medicine, or other. Respondents also indicated their year of graduation from medical school, the city in which they worked, and the type of hospital in which they worked (children’s hospital, university hospital, “community” hospital, or other). Twenty-three respondents identified themselves as pediatric infectious disease specialists and fifty-four identified themselves as pediatric emergency room specialists. The questionnaire presented the scenario of a child presenting to their emergency room with symptoms consistent with bacterial meningitis. Two related questions were asked: 1) “What do you think is the median time which elapses between the presentation of such a patient to the emergency room and the administration of intravenous antibiotics?”; and 2) “What length of time — from initial presentation to the E.R. until I.V. antibiotic administration — would encompass 95% of all such cases; i.e. of all such patients presenting to the E.R., what time would demarcate the point where 95% of children have received I.V. antibiotics?” William Meadow, Questionnaire, 1990 (on file with the *Duke Law Journal*).

such cases in their own hospitals. We then compared the responses of these doctors to the actual time to antibiotic administration (ABTIME) observed in ninety-three children treated at two large university-associated pediatric centers in Chicago.

We found, as we had hypothesized, that the recollections of the potential expert witnesses were wrong. Moreover, they were wrong in the anticipated direction—that is, biased in favor of the response that, in retrospect, would be desired (in this case, toward a shorter ABTIME). The median estimate of the fifty-four ER specialists was 46 minutes; the median estimate of the twenty-three ID specialists was 80 minutes. By contrast, the median ABTIME of ninety-three actual cases at the two university centers was 120 minutes, significantly longer than the estimates of either of the sub-specialty groups.³⁴ When we reviewed the medical literature in an attempt to extend these observations beyond Chicago, we found two hundred reported cases of children with meningitis from hospitals in South Carolina³⁵ and California.³⁶ These two hundred cases had a median ABTIME (114–126 minutes) comparable to the ninety-three Chicago cases, and longer than the estimates of either of the two sub-specialties.³⁷

34. Patient data were compiled from both hospitals independently. ABTIME values were not normally distributed for either institution. Consequently, the comparison of ABTIME between hospitals was performed nonparametrically using the Mann-Whitney Rank Sum test. As no significant difference in ABTIME was noted comparing the two hospitals, data from all meningitis patients were combined for subsequent analyses. The responses of the ID vs. ER experts differed significantly from each other (median ER ABTIME = 46 mins; median ID ABTIME = 80 mins; $Z_t = 2.28$; $p = 0.023$; ER vs. ID, Mann-Whitney Rank Sum Test), and both were significantly shorter than the actual median value of ABTIME (120 minutes) noted in our study population ($H = 45.33$, $p < 0.001$, Kruskal-Wallis Test; $Z_t = 2.67$, $p < 0.008$, ABTIME vs. ID ABTIME, Mann-Whitney Rank Sum Test; $Z_t = 6.60$, $p < 0.001$, ABTIME vs. ER ABTIME, Mann-Whitney Rank Sum Test) (aggregate results on file with the *Duke Law Journal*).

35. Charles S. Bryan et al., *Promptness of Antibiotic Therapy in Acute Bacterial Meningitis*, 15 ANNALS EMERGENCY MED. 544, 545 (1986).

36. David A. Talan et al., *Analysis of Emergency Department Management of Suspected Bacterial Meningitis*, 18 ANNALS EMERGENCY MED. 856, 856 (1989).

37. The university hospitals in Chicago are representative of what might be thought of as a national norm for the best possible treatment situation for a child with meningitis. Necessary equipment (i.v. catheters, antibiotic solutions) and experienced personnel (triage nurses, emergency physicians) are immediately available for children with meningitis at academic medical centers. Some have expressed concerns that for many cases this is too high a bar. See Sydney S. Gellis, "Standard Care" vs. "Standard of Care" in Administration of Antibiotics in Children with Meningitis, PEDIATRIC NOTES, Feb. 18, 1993, at 25:

Each infant or child presenting to an ER with possible meningitis is an individual problem and one which cannot be measured against others. . . . Timing will depend on the numbers of physicians available, numbers of critically ill patients being seen at the time, the availability of nursing help, [and] the presence of language barriers

Table 1 captures the differences between the opinions of potential experts and the actual data-based observations:

TABLE 1: VARIATIONS BETWEEN "EXPERT" OPINIONS AND DATA-BASED OBSERVATIONS FOR ABTIME IN CHILDREN WITH SUSPECTED BACTERIAL MENINGITIS

ID Experts	ER Experts	Chicago Hospitals	SC and CA Hospitals
80 minutes*	46 minutes**	120 Minutes	114–126 minutes

* $p < 0.01$

** $p < 0.001$

Whatever "expert" opinion means in this context, it does not mean an accurate opinion. Both groups of potential pediatric experts were wrong, and substantially so, in their opinions about the amount of time that elapses before doctors administer antibiotics to children with meningitis. The statistical probability that the potential experts were correct was less than one part in 100 for the ID doctors, and less than one part in 1000 for the ER doctors. As we have emphasized, the inaccuracy found in the estimates was consistent with our hypothesis: responses were biased toward the outcome perceived to be desired, that is, toward shorter waiting periods.

Note that this bias operates independently of the pressures imposed by the adversary system, which would naturally lead litigants to hire experts whose accounts will favor their side. The ID and ER respondents were not asked to consider their responses in the context of a review for either a plaintiff or defendant in this survey. Moreover, the optimistic bias presented by the potential experts here is operating outside of the bounds of the usual doctor-patient relationship, where it has been suggested that optimism, even if inaccurate, may serve a therapeutic purpose.³⁸

We think it entirely possible that other circumstances (e.g., smaller hospitals, clinics, or doctors' offices) may not be comparable to academic medical centers, and that ABTIME in these settings consequently might be longer than at university centers. We would then argue that claims about care in those smaller centers ought to be met with ABTIME data gathered from these clinics or offices. Our main point here is that if the standard of care is normative, appropriate data-based comparisons are best for determining what is normative.

38. See *supra* note 17 and accompanying text. There are complex relations between the malpractice setting and other contexts involving negligence. For example, the question, "How difficult would it have been for the worker to have avoided this hazard?" might lead to the (optimistic) suggestion that it would not have been difficult at all, whereas the question, "How ex-

Others, in varying contexts, have demonstrated a similar disconnect between physicians' descriptions of their own practices and objective determinations of the same behavior. As examples, consider patient education by physicians in general practice. Physicians are committed to the value of patient education,³⁹ and they devote considerable time and resources to health promotion.⁴⁰ Clearly, the desired practice is to spend this time explicitly informing patients about health-related behaviors, such as the uncontested virtues of smoking cessation and preventive oncology. Nevertheless, when physicians' reports of their office practices are compared with taped interviews of these same encounters, there is only a weak, insignificant correlation between the self-reported patient education activities and the actual performance.⁴¹ Not surprisingly, physicians' recollections err in the predicted direction; that is, more explicit counseling is recalled than actually occurred.

Here too the cognitive expectation is confirmed—anecdotal recall proves to be inaccurate, biased, and, inevitably, systematically imperfect. If our findings here are the rule and not the exception, as the psychological evidence suggests, it is hardly clear that the law should continue to base a system of medical-legal jurisprudence on such a shaky foundation.

pensive would it have been to have installed a safety device?" might lead to the (optimistic) suggestion that this would not have been expensive at all. Hence the framing of the question might lead optimistically biased agents in different directions. In either case, statistics are better than expert testimony, simply because they are more accurate.

39. Thomas E. Kottke et al., *Attributes of Successful Smoking Cessation Interventions in Medical Practice*, 259 JAMA 2883, 2887 (1988) (finding face-to-face advice from physician and nonphysician counselors to be the best predictor of smoking cessation); Henry Wechsler et al., *The Physician's Role in Health Promotion—A Survey of Primary-Care Practitioners*, 308 NEW ENG. J. MED. 97, 98–99 (1983) (finding that nearly three of four physicians feel it is a physician's responsibility to educate patients on risk factors including smoking, alcohol, drugs, stress, exercise, and diet, and that eighty-one percent of physicians interviewed personally provide such education in place of a nurse or other health professional).

40. C. Tracy Orleans et al., *Health Promotion in Primary Care: A Survey of U.S. Family Practitioners*, 14 PREVENTIVE MED. 636, 643 (1985) (surveying 610 family practice physicians on client demographics and medical treatment); see also Ulrich J. Grueninger et al., *Patient Education in the Medical Encounter: How to Facilitate Learning, Behavior Change, and Coping*, in THE MEDICAL INTERVIEW 122, 122–33 (Mack Lipkin, Jr. ed., 1995) (encouraging physicians not to simply provide information, but to incorporate educational and behavioral techniques).

41. Orleans et al., *supra* note 40, at 643; see also James E. Davis et al., *Cancer Prevention and Screening Activities in Primary Care Practice*, 16 PREVENTIVE MED. 277, 277 (1987) (comparing estimates by Wisconsin physicians of preventive performance with their actual preventive practices).

B. *The Proposal*

We acknowledge that the evidence we have offered is suggestive rather than demonstrative. Perhaps there are contexts in which doctors, or other experts, have an accurate sense of the ordinary standard of care. But the general evidence of mistakes, together with the empirical evidence introduced here, is sufficient to show that both error and optimism are highly likely to infect a wide range of expert testimony.⁴² This point operates independently of the ordinary incentive to assist one's side in an adversary system.⁴³

It therefore makes sense, in the context of malpractice suits and probably more generally as well, to move toward greater reliance on actual data and less reliance on the recollections of isolated experts. Until recently, the legal system was unable to rely on statistical data for the simple reason that it did not exist. But it is now increasingly possible to develop data sets about physician choices and behavior, and the legal system will have an increasing amount of information on which to draw. Perhaps government should generate accounts of ordinary practice, based on the information that is available; or perhaps people in the private sector should do so. In either case, the goal would be to ensure the compilation of actual data about practice, substituting accurate empirical evidence for the fallible judgments of individuals.

Our proposal would have two notable advantages. First, it could simplify the task of discerning the truth. It is now exceedingly difficult for juries to decide which expert has best described the prevailing standard of medical care in the community. Under current practice, this issue may well turn on unreliable judgments about credibility or on sympathy that is unrelated to the standard of care. Second, and more important, use of statistical evidence would increase the accuracy of assessments of the standard of care, reflecting the comparative advantage of pooled data over individual recollections. There is no good reason for courts to continue to rely on the latter when they have access to the former.

42. See Kahneman & Lovallo, *supra* note 18, at 18–21, 24–27 (discussing individuals' erroneous and overly optimistic decision-framing of probabilities).

43. See Linda Babcock & George Loewenstein, *Explaining Bargaining Impasse: The Role of Self-Serving Biases*, in *BEHAVIORAL LAW AND ECONOMICS*, *supra* note 8, at 355, 355–65 (describing self-serving bias in judgments about fairness).

III. PROBLEMS AND COUNTERARGUMENTS

A. *Statistical Error?*

Our proposal might be criticized on the ground that statistical analysis is itself subject to error. Of course there can be no assurance that any particular compilation or interpretation of data is correct. Perhaps there will be several accounts of data that will compete with one another. If this is so, it might be objected, the use of statistical evidence will merely usher in a new phase in the battle between imperfect experts. How much would be gained by that?

This objection has a degree of plausibility, but rather than impeaching our proposal, it suggests the need to develop good methods to evaluate particular claims about the meaning of statistical data. In the event of disagreement, there is no escaping the need to test the data in the ordinary way through the presentation of conflicting views, including those of experts. Data might be met with contrary data; it also might be met with a professional critique. And of course it is possible that an individual expert will be able to show persuasively that some purported evidence does not establish what it claims to establish. These sorts of disputes can be handled in the standard fashion. We suggest that because individual experts are prone to systematic error, it would be far better to begin the process with reliable evidence rather than with individual recollection. If individual experts can show that the statistical data are wrong, the legal system will be better off for the demonstration. In the long run, however, we predict that these demonstrations will be the exception rather than the rule. In any case, an inquiry into the evidence will anchor the legal process in something that is likely to be more reliable.

B. *Admissibility: Relevance and Hearsay*

Statistical data are not typically used in negligence cases. The data might be excluded as inadmissible on several grounds. Most important, questions might be raised about whether the data are relevant and about whether they are hearsay.

It is possible to object that statistical data are not relevant, because they cannot determine the proper standard of care. In some states, there is a continuing dispute about the proper role of common

practice in negligence cases.⁴⁴ Where common practice is not determinative of the standard of care,⁴⁵ statistical data cannot resolve the legal issue. It would be possible to argue that in such circumstances, the data should be inadmissible on grounds of irrelevance. But this argument should be rejected. Even when common practice is not determinative, it is pertinent, and its relevance is reflected by the very fact that experts frequently are allowed to testify about what most doctors do. Statistical data should be found admissible for the same reason that expert testimony is admissible.

It also might be thought that statistical data should be seen as hearsay. Perhaps the witness seeking to introduce the data is attempting to repeat what others have said in their reports of the data. If so, his testimony might be thought to violate the prohibition on hearsay.⁴⁶ But this analysis would misread the hearsay rules. Under federal law, for example, experts are allowed to rely on "statements contained in published treatises, periodicals, or pamphlets on a subject of history, medicine, or other science or art, established as a reliable authority by the testimony or admission of the witness or by other expert testimony or by judicial notice."⁴⁷ Federal rules also allow experts to base opinions on facts and data of the kind reasonably used by experts in the field when they form opinions or inferences.⁴⁸ Statistical data, if published, should be permitted under the first rule just quoted; if unpublished, they should be admissible to the extent that they are of the sort reasonably relied on by experts. Under either scenario, however, data should not be treated as inadmissible hearsay.

44. See *supra* note 1; *infra* note 46.

45. See, e.g., *Advincula v. United Blood Servs.*, 678 N.E.2d 1009, 1027–28 (Ill. 1996) (finding that evidence of compliance by the defendant blood bank with the professional standard of care imposed under a blood shield statute was not conclusive proof of "due care"); *Roach v. Springfield Clinic*, No. 73,394, 1992 Ill. LEXIS 204, at *33–*36 (Ill. Dec. 4, 1992) (finding that the exclusion of evidence describing medical standards of care did not prejudice a malpractice trial outcome); *Darling v. Charleston Cnty. Mem'l Hosp.*, 211 N.E.2d 253, 258 (Ill. 1965) (finding that the hospital's standard practice and conformity to regulations did not constitute conclusive evidence of reasonable care).

46. See *People v. Anderson*, 495 N.E.2d 485, 487–90 (Ill. 1986) (permitting a psychiatric expert to disclose the facts and opinions from others' evaluations on which he relied in diagnosis); *Schrag v. Chi. City Ry. Co.*, 106 N.E. 828, 829–30 (Ill. 1914) (prohibiting the admission of scientific books into evidence, as well as their use to contradict an expert witness not relying upon them).

47. FED. R. EVID. 803(18).

48. FED. R. EVID. 703, 705. But see *Roach*, 1992 Ill. LEXIS, at *32–*33 (refusing to allow references to treatises on the ground that their authors are not available for cross-examination). The *Roach* decision seems to us an anachronistic and odd reading of the prohibition on hearsay.

C. *Textbooks and Treatises*

Perhaps there are other solutions, besides reliance on statistical data, to the problem of idiosyncratic recollections of individual experts. It might be urged that the opinions of experts, as reported in textbooks or journal publications, should define the standard of care. By this view, when actual practices differ from the theoretical or recommended standards, the practitioners are at fault. They may not know the standard, may disagree with it, or may have misinterpreted it, but the burden of proof falls on them.

We believe that this approach would be oversimplified and misguided. Under the law, ordinary practice is important for its own sake, regardless of whether it is decisive of the legally acceptable standard of care. If courts are going to consider the ordinary practice, they should have an accurate understanding of what it is; here, textbook and treatise authors, like all other experts, are prone to error. Unless a textbook or treatise actually draws its conclusions from ordinary practice, it will not report the ordinary practice reliably.

Perhaps textbooks and journals merely state recommendations that are worth consideration even if they do not capture ordinary practice. This is not implausible. But interpreting these recommended standards is not always straightforward. In the particular context of bacterial meningitis, textbooks recommend that antibiotics should be administered "early,"⁴⁹ "promptly,"⁵⁰ or "immediately"⁵¹ to children with suspected meningitis. But what, in the minds of a lay jury, is a reasonable interpretation of this recommendation? In such a context, is thirty minutes too long? Is one hour, or six hours? The recommendation in one text that antibiotics be administered "within 30 minutes after the diagnosis of meningitis is established"⁵² helps a good deal, but it does not resolve the question, as it merely presents another troubling question—when exactly is the "diagnosis of meningitis established?"

But there is a more general point. At most, the author of a textbook can offer an opinion about the standard practice or about the

49. ABRAHAM M. RUDOLPH, *PEDIATRICS* 422 (16th ed. 1977).

50. 1 RALPH D. FEIGIN & JAMES D. CHERRY, *TEXTBOOK OF PEDIATRIC INFECTIOUS DISEASES* 302 (1981).

51. RICHARD E. BEHRMAN & VICTOR C. VAUGHAN, *TEXTBOOK OF PEDIATRICS* 571 (Waldo E. Nelson ed., 13th ed. 1987).

52. GERALD L. MANDELL ET AL., *PRINCIPLES AND PRACTICE OF INFECTIOUS DISEASES* 738 (1979).

practice that is desirable. If the opinion involves what is desirable, it should be evaluated as such. But if the opinion involves what is standard, the question should be whether the opinion is correct. To the extent that the question inquires into ordinary practice—into what doctors do—it would be far better to rely on actual data about physician performance.

Our plea here has been for attention to data, rather than to opinions. Ultimately we might hope for authoritative texts on what is established by the data, and such texts might come from the government or from the private sector. To the extent that there is competition among sources of data, all the better. And to the extent that data about actual practice are available, the data should inform, and improve, the debate about the extent to which the legal system should be bound by the *ordinary* standard of care when determining the *appropriate* standard of care.⁵³

CONCLUSION

It is well known that ordinary people make systematic errors in assessing probabilities and risks.⁵⁴ It is less well known that experts are subject to the same biases and likely to make the same mistakes.⁵⁵ We have suggested here that there is general reason to believe that expert judgments about the standard of medical care will be erroneous, and that the errors will run in a predictable direction because of optimistic bias. People tend to believe that things can be done, and are in fact done, more easily, more rapidly, and more successfully than the evidence suggests. To establish this claim, we have drawn on existing evidence, highly suggestive on this point, and on more particular evidence, presented here, that experts do not accurately report the ordinary standard of care.

Our principal innovation has been to suggest that in light of the evident divergence between expert beliefs and empirical reality, the legal system should rely, whenever possible, not on the former but on

53. See *supra* note 3 and accompanying text (discussing the relevance of customary practice in negligence judgments).

54. W. KIP VISCUSI, RATIONAL RISK POLICY 5 (1998); Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, in JUDGMENT AND DECISION MAKING 38 (Terry Connolly et al. eds., 2d ed. 2000); W. Kip Viscusi, *Jurors, Judges, and the Mistreatment of Risk by the Courts*, 30 J. LEGAL STUD. 107, 107 (2001).

55. See *supra* note 42 and accompanying text (discussing the inaccuracy of expert judgment that results from error and optimism).

statistical evidence of the latter. The best reason for the legal system's longstanding reliance on individual recollections has been historical—the simple absence of statistical evidence. But the current practice is an anachronism. The reason is that this gap in statistical evidence is rapidly being filled with a great deal of reliable data, and it may be completely filled in the next generation. Our emphasis has been on the question of medical malpractice, but the general implication is far broader. In any case in which a disputed question calls for expert testimony about ordinary practice, it is hazardous to rely on what particular experts recall. If the goal is accuracy in adjudication or regulation, it is far more sensible to make the outcome turn on statistical evidence.