THE SWINE FLU VACCINE AND GUILLAIN-BARRÉ SYNDROME: A CASE STUDY IN RELATIVE RISK AND SPECIFIC CAUSATION

by D.A. Freedman and P.B. Stark
Statistics Department, U.C. Berkeley, CA 94720

Technical Report No. 546 15 August 1999

Prepared for *Evaluation Review*

ABSTRACT

Epidemiologic methods were developed to prove general causation: identifying exposures that increase the risk of particular diseases. Courts often are more interested in specific causation: on balance of probabilities, was the plaintiff's disease caused by exposure to the agent in question? Some authorities have suggested that a relative risk greater than 2.0 meets the standard of proof for specific causation. Such a definite criterion is appealing, but there are difficulties. Bias and confounding are familiar problems; individual differences must be considered too. The issues are explored in the context of the swine flu vaccine and Guillain-Barré syndrome. The conclusion: there is a considerable gap between relative risks and proof of specific causation.

1. INTRODUCTION

In a toxic tort case, the plaintiff is exposed to a toxic agent, suffers injury, and sues. To win, the plaintiff must prove (i) "general causation" (the agent is capable of producing the type of injury in question), and (ii) "specific causation" (plaintiff's particular injury did in fact result from exposure to the agent). Both elements must be proved by the preponderance of the evidence. This standard is also characterized as "balance of probabilities," or "more likely true than not." Courts have used relative risk to assess evidence, with a relative risk above 2.0 arguing for specific causation. There are comments from various perspectives in Black and Lilienfeld (1984), American Medical Association (1987), Bailey, Gordis, and Green (1994), or Petitti (1996).

The intuition can be expressed as follows. Suppose that the exposed and unexposed groups in an epidemiologic study have the same distribution of variables affecting likelihood of injury—except for the particular exposure of interest. For simplicity, suppose the two groups are the same size. To have specific numbers, suppose there are 400 injuries among the exposed and only 100 among the unexposed, so the relative risk is 4. The conclusion is this: but for the exposure, there would be only 100 injuries among the exposed instead of 400. In other words, 300 of the 400 injuries are attributable to the exposure and 100 to other factors. Therefore, an injury among the exposed has chance 3/4 of being attributable to exposure. Likewise, a relative risk of 3 corresponds to a chance of 2/3, while a relative risk of 2 corresponds to a chance of 1/2, which is the breakpoint.

Here, we explore the scientific logic behind these intuitions. Of course, any epidemiologic study is likely to have problems of bias: uncontrolled confounding seems to be the rule rather than the exception (Freedman 1999, with citations to the literature). When effects are large, such problems may not be material; when relative risk is near the critical value of 2.0, potential biases need to be assessed more carefully. That is a serious and

generic problem, which we do not pursue. Individual differences also play an important role: plaintiff may not resemble typical members of the study population, and the effects of such differences need to be considered. Even in a randomized experiment, treatment and control groups are balanced in the aggregate but not at the level of individuals. This turns out to be a salient problem in connecting relative risk to specific causation.

We wanted to consider such issues in the context of a real example, in part to see how well the courtroom evidence stands up when examined retrospectively. Mike Green kindly provided a list of legal opinions where relative risk and specific causation come together (personal communication); also see Bailey, Gordis, and Green (1994, note 128). Generally, the underlying epidemiologic evidence was shaky. In one case—Manko v. United States—there turned out to be a substantial body of epidemiologic evidence, showing that the swine flu vaccine caused Guillain-Barré syndrome. And the vaccine campaign of 1976 is itself a fascinating case study. For these reasons, we selected *Manko* as our example; below, "the opinion" is the trial court's published decision, 636 F. Supp. 1419 (W.D. Mo. 1986). ("F. Supp." is the *Federal Supplement*; the opinion starts at page 1419 of volume 636; the case was heard in the federal district court for the Western District of Missouri.) The case was appealed, and the district court's opinion was affirmed in part, 830 F.2d 831 (8th Cir. 1987).

In section 2, we summarize the epidemiology of the swine flu vaccine and GBS (Guillain-Barré syndrome). GBS is a rare neurological disorder, sometimes triggered by vaccination or by infection; paralysis is a sequel, although most patients make a complete recovery in a few weeks or months. Section 3 turns to the Manko case and use of relative risk to demonstrate specific causation. Although the plaintiff prevailed, his proof of specific causation seems problematic to us, due in part to differences between him and typical members of the study population.

Appendix 1 presents a simple probability model where intuitions about relative risk and causation can be explored. The model sets aside all problems of confounding and bias, and considers only difficulties created by individual differences. For any particular plaintiff, the probability of causation is not identifiable from the data. Even the average probability of causation can be much lower than intuition suggests. For instance, if 4% of the exposed group suffers injury compared to 1% among the unexposed, the relative risk is 4 but the average probability of causation—given exposure and injury—can be as low as 3%, the difference in injury rates. We conclude that specific causation is linked to relative risk only through the principle of insufficient reason: the plaintiff is assumed to be like a randomly chosen subject in the epidemiologic study population, conditioned on exposure and injury.

2. THE SWINE FLU VACCINE AND GBS

In this section, we review the swine flu vaccination campaign of 1976 and the epidemiology of Guillain-Barré syndrome, as background for the discussion of *Manko* in Section 3. The influenza pandemic of 1918 killed some 20 million people world-wide. In February of 1976, a soldier in training at Fort Dix, New Jersey, died of influenza; the virus turned out to be similar in antigenic type to the 1918 virus. With public health professionals at the CDC (Centers for Disease Control) taking the lead, the Federal Government organized

a massive immunization campaign. Vaccination began on October 1, 1976. The vaccine was targeted at the 151 million people age 18 and over; some 43 million were eventually vaccinated. However, beyond the initial cluster at Fort Dix, only a handful of additional cases materialized, and several public health figures wanted the campaign stopped. A moratorium was declared on December 16, 1976—in part because an epidemic seemed increasingly unlikely, and in part because there were sporadic reports of GBS cases following vaccination. There are two different accounts of the vaccine campaign, by Neustadt and Fineberg (1982) and by Silverstein (1981). The latter was written to correct the former; but from our perspective, there is broad agreement on the central points.

The CDC played an active role in tracking reports of GBS and studying its relationship to the vaccine; the key papers are Langmuir (1979), Schonberger et al. (1979), and Langmuir et al. (1984); the last is perhaps the best single reference. The study period in these papers runs from October 1, 1976 to January 31, 1977. Langmuir et al. analyzed the incidence rate of GBS among the vaccinated, by weeks since vaccination; this rate is shown as the highly peaked solid line in Figure 1a (computed by us from their data). Rates are "per million person-weeks" of exposure; these are incidence rates, not relative risks. There is a strong peak for a few weeks after vaccination. In other words, there is a clear association between vaccination and GBS, provided the onset of GBS is within a few weeks of vaccination. Shown for comparison is the incidence rate of GBS among the unvaccinated, by calendar week from October 1st (lower dashed line, also computed from Langmuir et al.'s data). Notice that two time scales are involved. The sizes of the vaccinated and unvaccinated populations are changing rapidly over time, due to the vaccination campaign; size is taken into account in computing the rates.

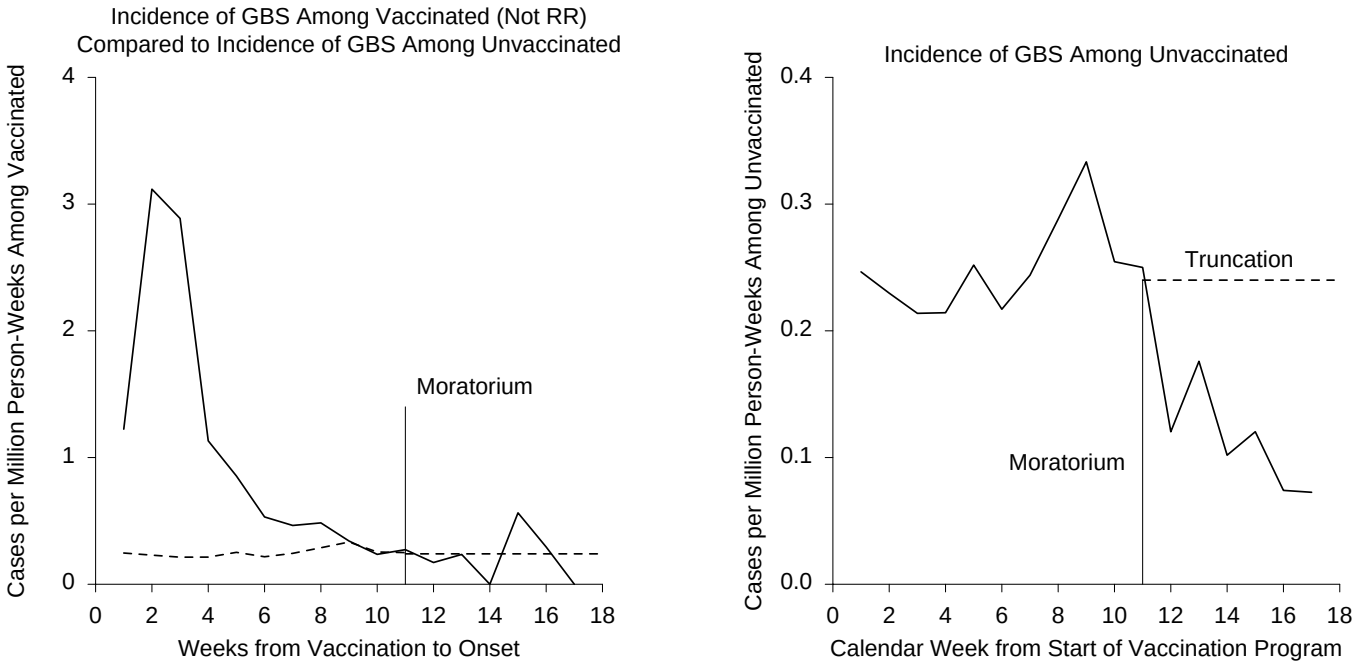
Is the association causal? That is still controversial. No excess risk for GBS was observed in the military, or with previous vaccines much like the swine flu vaccine. See, for instance, Hahn (1998, 636), Ropper et al. (1991, 28–9, 67), Hughes (1990, 103), Safranek et al. (1991), Beghi et al. (1985), Kurland et al. (1985, 636); Hughes and Rees (1997) find the evidence less ambiguous. In subsequent mass vaccinations, excess risk is minimal, although statistical significance is achieved if data for 1992–93 and 1993–94 are pooled (Lasky et al. 1998, Kaplan et al. 1982, Hurwitz et al. 1981). Further arguments and counter-arguments will not be discussed here. After reviewing the data and the literature, we think that a finding of general causation is reasonable: on balance of evidence, the swine flu vaccine did increase the risk for GBS for a period of several weeks after vaccination.

The background rate in Figure 1a is shown on a magnified scale in Figure 1b. After the moratorium, there is a precipitous drop in the background rate. This drop is best explained as an artifact of data collection. After the moratorium, it seems probable that GBS was less in the news, neurologists were less likely to make the diagnosis among unvaccinated persons, and state health departments were less diligent in collecting the data and reporting to CDC. Some of the drop may also be due to lags in data collection. Larry Schonberger, who was doing surveillance at the CDC, reports that a number of states put significantly less effort into data collection after the moratorium (personal communication; also see Schonberger et al. 1979, 197).

The background rate of GBS (among unvaccinated persons) is a critical baseline statistic: the incidence rate of GBS among the vaccinated persons is compared to this baseline,

in Figure 1 and in computations of relative risk. GBS is not a reportable disease, nor is the diagnosis easy. Thus, considerable uncertainty attaches to the background rate. Langmuir et al. (1984, 856) did not believe the background could be below 0.24 per million person-weeks, that is, about 1 case per 100,000 persons per year (for details, see Appendix 2). Thus, Figure 1a takes the background rate as 0.24 after the moratorium: the lower dashed line is horizontal after week 11. Current literature suggests a background rate of 0.2 to 0.4 per million person-weeks: see Hahn (1998, 635), Ropper et al. (1991, 19) or Hughes (1990, 101); but also see Lasky et al. (1998), who found a rate of about 0.15 per million person-weeks. So far as is known, there is no seasonal pattern to incidence rates of GBS.

FIGURE 1: Panel (a) shows the incidence rate among the vaccinated, by week since vaccination (highly peaked solid line); this rate is compared to the background rate (lower dashed line) among the unvaccinated, by week since the start of vaccination campaign. Two time scales are involved. The moratorium occurred in the 11th week after the start of the campaign, indicated by a vertical line. Panel (b) shows the background rate in more detail.



Other features of the data analysis in Langmuir et al. (1984) may be of interest. (i) They distinguish between cases with "extensive" and "limited" paralysis; the association is strong for the extensive cases, but there is little evidence of association for the limited cases. (ii) They fit log normal curves to the incidence data, and argue causation from the goodness of fit—which in retrospect seems curious.

A little more background is needed. Before the 1976 swine flu campaign got under way, the insurance companies refused to issue coverage for adverse events resulting from vaccination, and the drug companies refused to produce the vaccine without coverage.

To resolve this impasse, the Federal Government accepted liability in the Federal Tort Claims Act, 28 U.S.C. §1346(b). Thus, GBS victims applied for compensation not to the vaccine providers but to the Federal Government. There were about 510 GBS victims among the vaccinated and 440 among the unvaccinated. These generated nearly 2,000 legal claims, one of which is the topic of the next section. Current legal procedures for handling vaccine-related injuries are discussed by Johnson, Drew, and Miletich (1998).

3. THE MANKO CASE

In *Manko*, plaintiff used relative risk to demonstrate specific causation. This case was well argued, with a solid basis in epidemiology. Still, we find the proof unconvincing. The evidence will be reviewed in some detail, so the strengths and weaknesses of the relative-risk approach can be seen. Louis Manko was vaccinated on October 20, 1976 and developed symptoms of "smoldering GBS" within a week or two. Around January 15, 1977, he was hospitalized with acute GBS. The Federal Government refused compensation, on the theory that his "smoldering GBS" was not GBS, and his acute GBS developed too long after he was vaccinated for causation to be probable. Manko sued and the court ruled in his favor, adopting two theories of specific causation. (i) If "smoldering GBS" is indeed GBS, then causation follows from the epidemiologic evidence reviewed in Section 2 above. (ii) If on the other hand plaintiff contracted GBS in mid-January of 1977, some 13 weeks after vaccination, specific causation still follows because the relative risk for such late-onset cases is well above the threshhold value of 2.0.

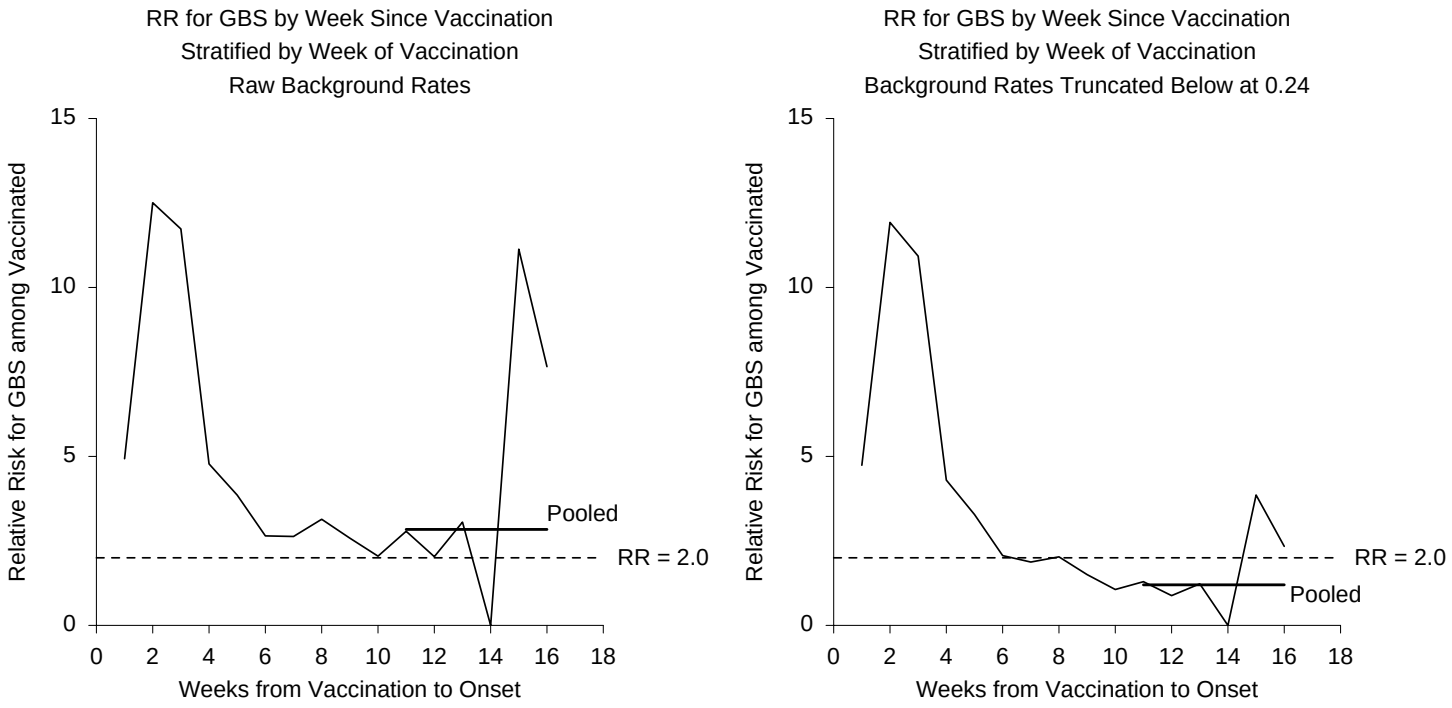
The arguments on causation for late-onset cases (pp. 1433ff in the opinion) are the most interesting. Plaintiff introduced expert testimony from Nathan Mantel and Martin Goldfield. Mantel was a well-known biostatistician at the National Institutes of Health. Goldfield was the county medical officer who worked on the Fort Dix outbreak; he was one of the first to identify the disease as influenza and one of the first to advise against mass vaccination. Defendants' epidemiology experts were Leonard Kurland of the Mayo Clinic and Neal Nathanson of the Pennsylvania Medical School; they were coauthors of the Langmuir report, cited here as Langmuir et al. (1984). Kurland and Nathanson are well respected in the field, and Langmuir was the founder of the Epidemiologic Investigative Service at the CDC.

Figure 1a—essentially the case for the defense on late-onset GBS cases—shows only a small excess risk after the 8th week. That figure stratifies on time since vaccination. However, Goldfield and Mantel argued that in order to compare like with like, it was also necessary to stratify on time of vaccination when computing relative risks. Their rationale was ingenious: they hypothesized a decrease in reporting of vaccinated GBS cases, parallel to the decline in reporting of the unvaccinated cases. Stratification is explained in Appendix 2, but the idea is simple. The relative risk compares the observed number of GBS cases with the number expected on the theory that vaccination does not cause GBS. Goldfield and Mantel computed the expected numbers separately for each vaccination cohort—those vaccinated in week 1, those vaccinated in week 2, and so forth. Finally, the contributions from the various cohorts are summed to get the expected number of cases in each week after vaccination. In effect, this synchronizes the two time scales in Figure 1a. However, the late-onset cases are being compared to the very small number of background

cases reported after the moratorium, so the estimated relative risk is large.

Goldfield and Mantel used the raw (untruncated) background rates to compute the relative risk, as in Figure 2a. For comparison, Figure 2b shows relative risks computed by the Goldfield-Mantel procedure, stratifying both on time of vaccination and time since vaccination, but with background rates truncated below at 0.24 per million person-weeks of exposure. The threshhold relative risk of 2.0 is marked by dashed horizontal lines. There were no cases in the 14th week after vaccination, only 4 in the 15th week, and 1 in the 16th week. The tail of the curve is quite shaky, so plaintiff's experts pooled the data for weeks 11–16 as indicated by the solid horizontal lines in both panels. Both panels in Figure 2 use the same observed numbers and compute expected numbers the same way—except for truncation. The issue is not the stratification but the truncation, and the crucial question is this: was there a drop in reporting of vaccinated GBS cases after the moratorium, parallel to the drop in background rates? If so, Figure 2a is persuasive and the relative risk for late-onset cases is high. If not, Figure 2b is the one to use and excess risk is minimal.

FIGURE 2: Panel (a) shows the Goldfield-Mantel analysis, with stratification by time of vaccination as well as time since vaccination; raw background rates are used. Panel (b) stratifies the same way, but truncates the background rate from below. The short horizontal line pools the data in weeks 11–16, to stabilize the estimates.



COMPLETENESS OF REPORTING

Both sides in *Manko* agreed that the drop in background rates was artifactual (Goldfield, Tr. 6.44, for the plaintiff; Langmuir et al. 1984, 856 for the defense; compare Tr. 18.114–5 and p. 1435 of the opinion; "Tr. 6.44" is p. 44 of vol. 6 of the trial transcript). The issue was the plaintiff's hypothesis of a parallel drop in reporting of vaccinated cases. To validate that hypothesis, Goldfield and Mantel (Tr. 6.61–67, esp. Tr. 6.66) compared the incidence rate of GBS among the vaccinated before and after the decline in background rates. However, the numbers are small. Furthermore, a real decline in the incidence rate is only to be expected, because the attack rate decreases with time since vaccination (Figure 1), and most vaccinations occurred fairly early in the sequence of events. Thus, it is not easy to show that decline in reported incidence rates is too abrupt, although Figure 1 in Schonberger et al. (1979) suggests there is something to the idea; also see Langmuir et al. (1984, Table 5).

To address the completeness of reporting, Langmuir et al. (1984, 860ff) compared attack rates for three cohorts—persons with early, middle, and late vaccinations. They saw no evidence for a decline in reporting rate among vaccinated GBS cases. Vaccination by itself could have made a diagnosis of GBS more likely, because vaccination was seen as a leading cause of GBS. Moreover, reporting is likely a priori to be more complete among the vaccinated cases than the unvaccinated: vaccinated cases had to be reported to the Federal Government in order for victims to claim compensation. As noted above, 510 vaccinated cases gave rise to 2,000 legal claims, which hardly suggests under-reporting. Also see Marks and Halpin (1980, 2493).

For an empirical test, we followed Goldfield and smoothed the relative risks in Figure 2b. We then used the smoothed curve to estimate the likely number of post-moratorium GBS cases among the vaccinated; details are in Appendix 2. Specific results depend on the smoothing. However, as best we can tell, the reporting of vaccinated GBS cases dropped after the moratorium by no more than 20%. There also seems to have been over-reporting in weeks 10 and 11: perhaps onset dates were advanced by a week or two in the CDC's database, around the time of the moratorium. Although the relative risk for late-onset GBS in Figure 2b may biased downward, the effect seems to be small. Indeed, current medical literature does not support the hypothesis of swine flu vaccination as a cause of late-onset GBS cases (Hahn 1998, 636; Ropper et al. 1991, 28–9, 57; Hughes 1990, 102).

DISCOVERY ISSUES

There is now another legal complication. In pre-trial discovery proceedings, each side gets to demand documents from the other. However, the Federal Government declined to produce the CDC's detailed medical records on GBS victims. In some of these cases, critical information on the date of vaccination or the date of onset of GBS was missing on the summary sheets that were made public and used both by plaintiffs and defense. To resolve this discovery issue, the court imposed a sanction. Langmuir et al. (1984, 845) had excluded from their analyses some 28 cases with missing data; plaintiff's experts were allowed to count 8 of these cases as having late onset. Table 1 shows the relative risk for GBS with onset 11–16 weeks after vaccination, computed on the various sets of assumptions. As Table 1 confirms, the critical issue is the truncation. (See pp. 1438 and

1453 of the opinion on sanctions, and pp. 1436–7 on the calculation of relative risk; we infer the figure of 8 additional cases to reconcile the numbers in notes 10 and 11 of the opinion with the data in Langmuir et al.)

TABLE 1: Relative risks for GBS cases, with onsets in weeks 11–16 after vaccination. RR = Observed/Expected. The first column computes the "Expected" with the background rate truncated below. The second column uses the raw background rate. Row 1 shows data for extensive cases; row 2, for all cases; row 3 adds 8 cases to the numerator, as a consequence of sanctions imposed by the court on defendants.

	Truncated below	Raw
Extensive cases	$9/10.2 = 0.88$	$9/4.41 = 2.04$
All cases	$21/17.5 = 1.20$	$21/7.40 = 2.84$
Sanctions	$29/17.5 = 1.66$	$29/7.40 = 3.92$

NOTE: The numbers in the table are computed by us from data in Langmuir et al. (1984); for details, see Appendix 2.

INDIVIDUAL DIFFERENCES

We turn now to individual differences. Prior infection is a risk factor for GBS: about 62% of the unvaccinated GBS cases had some illness in the month before onset. For the vaccinated cases, only 33% had prior illness. See Schonberger et al. (1979, 116) or Langmuir (1979, 663). An informal calculation (Appendix 2) suggests that prior illness multiplies the relative risk by about $33/62 \doteq 0.53$. Manko had an infection with respiratory and gastrointestinal symptoms a week or two before his hospitalization for acute GBS (plaintiff's exhibit 401): multiplying the relative risk of 3.92 by .53 brings it very close to the critical value of 2.0. However, Goldfield and Mantel argued that the .53 includes a selection effect: people are advised against vaccination immediately following illness. To avoid the selection effect, Goldfield and Mantel based their correction only on the late-onset GBS cases among vaccinated persons, where 53% were preceded by illness (Tr. 7.39); the relative risk should now be multiplied by $53/62 \doteq 0.85$. (In note 12 to the opinion, the multiplier is given as 0.87; different experts—even on the same side—seem to have been using slightly different versions of the CDC database; and there is an annoying numerical coincidence, as .53 crops up twice with two different meanings.)

The number of late onset cases is rather small (Table 1), and the experience of this group should probably not be compared to all unvaccinated cases but to cases with onsets in a similar time period—late December and January—because the pattern of background illness is quite seasonal. Plaintiff's argument is therefore not wholly convincing. Current literature confirms that about 2/3 of GBS cases are triggered by previous illness; see, for instance, Hahn (1998, 636), Ropper et al. (1991, 57), or Hughes (1990, 106). With

respect to one pathogen, *Campylobacter jejuni*—which causes gastrointestinal symptoms—the molecular basis for subsequent GBS is now reasonably well understood (Nachamkin et al. 1998).

Age is another factor to consider. Manko was 64 years old at vaccination (plaintiff's exhibit 401, Tr. 16.193). That would reduce the relative risk by perhaps 25%, if it is fair to average across onset times (Schonberger et al. 1979, 114; Lasky et al. 1998, Table 1). Finally, the clinical course of the disease should be mentioned. About 95% of patients reach their nadir within a month of onset, and roughly 70% recover completely within a year (Hahn 1998, 639; Hughes 1990, 122–23; compare p. 1427 of the opinion). In this respect too, Manko was quite unlike the bulk of the GBS victims. Therefore, the data cannot tell us very much about the cause of Manko's injury. *Manko* was a well-argued case with a solid empirical base reported in the epidemiologic literature. Even so, the proof of specific causation—starting from a relative risk of 4—seems unconvincing. That gives us pause, and the issue goes well beyond *Manko*. The impact of individual differences on the probability of specific causation is discussed analytically, in Appendix 1.

4. SUMMARY AND CONCLUSIONS

The scientific connection between a relative risk of 2.0 and specific causation is doubtful. Large relative risks argue for general causation, while small ones argue against. If the relative risk is near 2.0, problems of bias and confounding in the underlying epidemiologic studies may be serious, perhaps intractable. Problems created by individual differences may be equally difficult. Bias and confounding affect the estimation of relative risk from the underlying data. By contrast, individual differences affect the interpretation of relative risk—namely, the application to any specific individual.

With *Manko*, it was difficult to establish an elevated relative risk for late-onset cases. Moreover, the plaintiff was in crucial detail remarkably unlike the other GBS victims. So the connection between him and the data stayed rather loose. The mathematical models developed in Appendix 1 show how the effect of individual differences can be represented in a more general—but more abstract—setting. The results confirm one of our central points about *Manko*: epidemiologic data cannot restrict the probability of causation to any useful degree, because of individual differences.

In law cases, the plaintiff always has the burden of persuasion; but the burden of production—of going forward with the argument—does shift. If introducing epidemiologic evidence on relative risk shifts the burden of production, that has consequences: quantifying the impact of deficiencies in a study can be even harder than quantifying the risk associated with exposure.

APPENDIX 1. A PROBABILITY MODEL

This appendix presents a probability model for the chance that exposure causes injury. The object of the model is to clarify the statistical problem, and show that the probability of causation cannot be determined from aggregate data, except within very broad limits. Our model is a variant of Neyman's counterfactual model for causal inference (Neyman 1923, translated by Dabrowska and Speed 1990; Hodges and Lehmann 1964; Rubin 1974; Holland 1988). Although counterfactuals may seem a bit academic, the legal idea of

causation is quite consistent with the model: but for exposure, plaintiff would not have been injured (Hart and Honoré 1985, 104). "Exposure" and "injury" are used generically. In the Manko case, for instance, exposure was to the swine flu vaccine and the injury was GBS; with smoking and lung cancer, exposure is to cigarette smoke and the injury is lung cancer.

In the model, there are n subjects. Associated with each subject i are three random variables, U_i, V_i, X_i. The random variable U_i is the response if subject i is not exposed: $U_i = 1$ if subject i would then be injured, and $U_i = 0$ if subject i would not be injured. The variable V_i represents subject i's response to exposure, defined analogously. The pairs (U_i, V_i) are assumed to be independent across i. In this model, both U_i and V_i exist, whether or not subject i is exposed.

Each subject is characterized by a 4-vector of probabilities, p_i, q_i, r_i, s_i, specifying the joint distribution of (U_i, V_i):

$$\begin{aligned} p_i &= P\{U_i = 0 \text{ and } V_i = 0\}, \\ q_i &= P\{U_i = 0 \text{ and } V_i = 1\}, \\ r_i &= P\{U_i = 1 \text{ and } V_i = 0\}, \\ s_i &= P\{U_i = 1 \text{ and } V_i = 1\}. \end{aligned}$$

Of course, p_i, q_i, r_i, and s_i are nonnegative and sum to 1.

The exposure variable X_i is defined to be 1 if subject i is exposed, and 0 otherwise. The epidemiologist observes U_i if $X_i = 0$, and V_i if $X_i = 1$. Let $\mathcal{X} = \{i : X_i = 1\}$ be the set of exposed subjects, and let $m = \sum_{i=1}^{n} X_i$ be the number of exposed subjects. Epidemiologic studies are generally observational, but to capture the idea of no confounding, we assume that nature has run a randomized experiment for the epidemiologist: (i) the exposure variables $X_1, \ldots, X_n$ are independent of the response variables $U_i, V_i : i = 1, \ldots, n$, and (ii) the exposure variables randomly select m of the n subjects to be exposed. In this setup, m is fixed, and $P\{X_i = 1\} = m/n$ for all i. We think of the model in a series of steps. For each i, the response variables U_i and V_i are drawn according to their joint distribution. Next, the exposure variables X_i are drawn, independently of the response variables U_i, V_i. The final step: if $X_i = 0$, then U_i is revealed; if on the other hand $X_i = 1$, then V_i is revealed.

The probabilities p_i, q_i, r_i, s_i are fixed and inherent characteristics of subject i. These probabilities generate U_i, V_i, which characterize subject i's response to non-exposure and exposure, respectively. Once generated, U_i and V_i also become fixed personal characteristics. Subjects operate independently of each other, in the sense that one person's response is unrelated to another's. After the response variables are determined, each subject's response is triggered by one thing only: whether the subject is or is not exposed. The random variables $X_1, \ldots, X_n$ describe the process by which subjects came to be exposed or unexposed. This process is assumed to be independent of the personal characteristics of the subjects, including their response variables.

Causation is defined in the following way. Suppose subject i is exposed and injured: $X_i = 1$ and $V_i = 1$. The injury is caused by the exposure if the injury would not have occurred but for the exposure, that is, $U_i = 0$. After all, if $U_i = V_i = 1$, then i would have

been injured whether exposed or unexposed. Said another way, the subject is known to have been exposed and injured. Causation means that if—counterfactually—the subject had not been exposed, the injury would not have occurred. We learned how to use counterfactual random variables to make this idea precise from James Robins (personal communication). Also see Robins and Greenland (1989ab).

The random variable X_i is observed: it is 1 if subject i is exposed, 0 otherwise. If $X_i = 1$, then V_i is observed; else, U_i is observed. Thus, exactly one of U_i, V_i is observed, and the other remains a counterfactual. Our condition for no confounding is stated in terms of both U_i and V_i; likewise, our theorems involve features of the joint distribution of U_i and V_i. In the legal setting, $X_i = 1$ and $V_i = 1$ as well: the plaintiff was exposed and injured. The argument is about U_i: would the plaintiff have been injured if not exposed? This question is inherently counterfactual; after data collection, U_i is the counterfactual remaining in the model to answer the question.

We turn now to the details. Let β_i be the chance that subject i would be injured if left unexposed,

$$\beta_i = P\{U_i = 1\} = r_i + s_i.$$

And let γ_i be the chance that subject i would be injured if exposed,

$$\gamma_i = P\{V_i = 1\} = q_i + s_i.$$

These probabilities are identifiable and estimable, although with poor precision: there is one observation for each person, and two parameters to estimate. The probabilities p_i, q_i, r_i and s_i are not separately identifiable. For instance, even if we knew that $r_i + s_i = .01$ and $q_i + s_i = .04$, we could not recover q_i without further information

Let β be the expected rate of injury if all subjects are unexposed:

$$\beta = \frac{1}{n}\sum_{i=1}^{n} P\{U_i = 1\} = \frac{1}{n}\sum_{i=1}^{n} (r_i + s_i). \tag{1}$$

Let γ be the expected rate of injury if all subjects are exposed:

$$\gamma = \frac{1}{n}\sum_{i=1}^{n} P\{V_i = 1\} = \frac{1}{n}\sum_{i=1}^{n} (q_i + s_i). \tag{2}$$

The difference $\gamma - \beta$ is the difference between the expected injury rates in two situations: (i) we expose everybody, and (ii) we expose nobody. In short, $\gamma - \beta$ is the average causal effect of exposure on injury, in our study population.

Given the assumption that the exposed group is a random sample of size m from the n subjects, there are unbiased estimators of β and γ:

$$\hat{\beta} = \frac{1}{n-m}\sum_{i \notin \mathcal{X}} U_i$$

and

$$\hat{\gamma} = \frac{1}{m} \sum_{i \in \mathcal{X}} V_i.$$

Unbiasedness is a consequence of the no-confounding assumption.

To avoid irrelevant technicalities, we will assume that β and γ are known, with

$$0 < \beta < \gamma, \tag{3a}$$

$$0 < \beta + \gamma < 1. \tag{3b}$$

For present purposes, the relative risk is RR $= \gamma/\beta$. Condition (3a) says that exposure increases the expected injury rate, so the RR exceeds 1. Condition (3b) obtains in the usual situations when the injury rate is not high, even in the exposed group; this condition will be relaxed, below.

In the litigation context, we would like to find the conditional probability that exposure caused injury to a subject, given that the subject was exposed and injured. Because the exposure variables are independent of the response variables,

$$P\{U_i = 0 | V_i = 1, X_i = 1\} = P\{U_i = 0 | V_i = 1\}. \tag{4}$$

The conditional probability that subject i's injury was caused by exposure is then

$$\pi_i = P\{U_i = 0 | V_i = 1\} = q_i/\gamma_i = q_i/(q_i + s_i). \tag{5}$$

The definition of π_i if $\gamma_i = 0$ is not material; for convenience, set $\pi_i = 0$ in this circumstance. Because q_i is not identifiable, neither is π_i.

Given that subject i was injured when exposed, π_i is the conditional probability that this subject would not have been injured if left unexposed: π_i is the probability of specific causation. If subject i brings suit, the court—in this formalism—may wish to decide whether $\pi_i > 1/2$; if so, specific causation is more probable than not. The question cannot be answered from epidemiologic data, because π_i is not identifiable. (The connection between evidence and probability might itself need clarification, but that is a topic for another essay.) Therefore, estimates for the probability of causation usually turn out to be based on assumptions that are largely untestable.

Our next result gives sharp bounds on $\overline{q} = \frac{1}{n} \sum_{i=1}^{n} q_i$. Theorem 1 below provides a sharp bound on $\overline{\pi}$, the average probability of causation across individuals. As will be seen, the bounds are usually quite broad: epidemiologic data cannot restrict the probability of causation for individuals in any meaningful way, or even the average probability across individuals.

Lemma 1. Fix β and γ satisfying (3). Suppose p_i, q_i, r_i, s_i are non-negative, sum to 1 for each i, and satisfy (1) and (2).

(a) $\overline{q} \leq \gamma$, with equality iff $s_i = 0$ for all i.
(b) $\overline{q} \geq \gamma - \beta$, with equality iff $r_i = 0$ for all i.

Proof. Claim (a) is obvious from (2); it is feasible to set $s \equiv 0$ by (3b). For claim (b), $\overline{q} = \gamma - \beta + \overline{r}$ by (1) and (2). The proof is complete.

The next theorem shows that the average probability of causation can be arbitrarily low, no matter how large the relative risk. It is the difference between risks that matters, not their ratio.

Theorem 1. Fix β and γ satisfying (3). Suppose p_i, q_i, r_i, s_i are non-negative, sum to 1 for each i, and satisfy (1) and (2). Define π_i by (5). Then $\inf \overline{\pi} = \gamma - \beta$, where the infimum is taken over n as well as the probabilities p_i, q_i, r_i, s_i. Requiring the probabilities to be strictly positive does not change the infimum.

Proof. To begin with, $\overline{\pi} \geq \gamma - \beta$:

$$\overline{\pi} = \frac{1}{n}\sum_{i=1}^{n} \pi_i \geq \frac{1}{n}\sum_{i=1}^{n}(q_i + s_i)\pi_i = \frac{1}{n}\sum_{i=1}^{n} q_i = \overline{q} \geq \gamma - \beta, \tag{6}$$

by Lemma 1b.

The lower bound is attained if $r_i = 0$ for all i, and $q_i = 0$ unless $q_i + s_i = 1$, which motivates the following construction. Suppose there are two types of subjects. With type A subjects, exposure does not change the probability of injury. Type B subjects, however, are injured only by exposure. More rigorously, suppose to begin with that β and γ are rational; choose n so that $n\beta$ and $n\gamma$ are integers. Divide $\{1, \ldots, n\}$ into two disjoint sets A, B with $|A| = n(1 - \gamma + \beta)$ and $|B| = n(\gamma - \beta)$, where $|C|$ is the cardinality of C. For $i \in A$, set

$$p_i = \frac{1-\gamma}{1-\gamma+\beta},$$

$q_i = r_i = 0$, and

$$s_i = \frac{\beta}{1-\gamma+\beta} = 1 - p_i.$$

For $i \in B$, set $p_i = 0$, $q_i = 1$, $r_i = s_i = 0$. Plainly, (1) and (2) hold, while $\overline{\pi} = \gamma - \beta$.

To make all probabilities positive, modify the construction as follows. Let $\delta > 0$ be small. For $i \in A$, set

$$p_i = \frac{1-\gamma}{1-\gamma+\beta} - \frac{1+\gamma-\beta}{1-\gamma+\beta}\,\delta,$$

$$q_i = \frac{1+3(\gamma-\beta)}{1-\gamma+\beta}\,\delta,$$

$$r_i = \delta,$$

and

$$s_i = \frac{\beta}{1-\gamma+\beta} - \frac{1+\gamma-\beta}{1-\gamma+\beta}\,\delta.$$

For $i \in B$, set $p_i = \delta$, $q_i = 1 - 3\delta$, $r_i = s_i = \delta$. For each i, the probabilities are positive and sum to 1. We have increased $\overline{r}$ by δ, decreased $\overline{s}$ by δ, and increased $\overline{q}$ by δ. Thus, (1) and (2) still hold, while $\overline{\pi} = \gamma - \beta + O(\delta)$.

To eliminate the restriction that β, γ are rational, fix $\epsilon > 0$. Let Δ_ϵ be the set of real β, γ for which $\epsilon \leq \beta$ and $\beta + \epsilon \leq \gamma$ and $\beta + \gamma \leq 1 - \epsilon$. This is a closed triangle. Let Δ'_ϵ be

the set of β, γ in Δ_ϵ for which the theorem holds: Δ'_ϵ is a closed subset of Δ_ϵ that contains all points with rational coordinates in Δ_ϵ. Therefore, $\Delta'_\epsilon = \Delta_\epsilon$. To complete the proof, let $\epsilon \to 0$.

Remark. The sup of $\overline{\pi}$ is easily seen to be 1: choose $s \equiv 0$, $r \equiv \beta$, $q \equiv \gamma$, and $p \equiv 1 - \beta - \gamma$, which is feasible by (3).

Terminology. The "inf" or "infimum" of a set S of numbers is the smallest element of S when that exists; otherwise, $\inf S$ is the largest number x with $x \leq s$ for all s in S. Likewise, the "sup" or "supremum" of S is the largest element of S when that exists; otherwise, $\sup S$ is the smallest number x with $x \geq s$ for all s in S. For instance, 0 is the inf of the positive real numbers, and the sup of the negative reals. By special convention, the sup of the positive real numbers is ∞, and the inf of the negative reals is $-\infty$. The "cardinality" of a set is the number of its elements: for instance, the cardinality of $\{A, B, C\}$ is 3.

Numerical examples. Suppose $\beta = .01$ and $\gamma = .04$, so the relative risk is 4.

(i) If $p_i = .95$, $q_i = .04$, $r_i = .01$, $s_i = 0$ then $\pi_i = 1$; this is possible for all i, because $\overline{p} = .95$, $\overline{q} = .04$, $\overline{r} = .01$, $\overline{s} = 0$, so $\beta = \overline{r} + \overline{s} = .01$ and $\gamma = \overline{q} + \overline{s} = .04$, as required.

(ii) If $p_i = .96$, $q_i = .03$, $r_i = 0$, $s_i = .01$ then $\pi_i = .75$; again, this is possible for all i, because β and γ have the right values.

(iii) If 97% of the subjects have

$$p_i = 96/97,\ q_i = r_i = 0,\ s_i = 1/97$$

and 3% have

$$p_i = 0,\ q_i = 1,\ r_i = s_i = 0,$$

then $\pi_i = 0$ for the first group and $\pi_i = 1$ for the second, so $\overline{\pi} = .03$, which is the lowest possible. As before, β and γ have the right values.

Notice that examples (ii) and (iii) have the same values for $\overline{p}$, $\overline{q}$, $\overline{r}$, $\overline{s}$ but radically different values for $\overline{\pi}$. Proponents of relative risk seem to be thinking of example (ii), but maybe they should also be thinking of examples (i) and (iii). Generally, epidemiologic studies cannot distinguish situation (ii) from (i) or (iii): only β and γ are determined.

Parametric models. The framework discussed here includes some familiar parametric models. For instance, consider the probit specification. Subject i has a latent variable W_i and a vector of personal characteristics Z_i. There is a causal parameter a and a vector b of nuisance parameters. It is assumed that the W's are independent of each other, the Z's, and the X's, and have a common normal distribution with mean 0 and variance 1. Let $U_i = 1$ if $Z_i b + W_i > 0$, otherwise $U_i = 0$. Similarly, let $V_i = 1$ if $a + Z_i b + W_i > 0$, otherwise $V_i = 0$. This leads to the probit model, which substantially restricts the inference problem, especially if Z_i is low-dimensional.

Remark. Suppose $0 < \beta, \gamma < 1$ but $\beta + \gamma \geq 1$. The lower bounds in Lemma 1 and Theorem 1 go through unchanged. However, the sharp upper bound on $\overline{q}$ is $1 - \beta$, achieved

when $\overline{s} = \beta + \gamma - 1$. Indeed, (1) and (2) entail

$$1 \geq \overline{q} + \overline{r} + \overline{s} = \beta + \gamma - \overline{s}$$

so

$$\overline{s} \geq \beta + \gamma - 1$$

and

$$\overline{q} = \gamma - \overline{s} \leq \gamma - (\beta + \gamma - 1) = 1 - \beta.$$

The sharp upper bound on $\overline{\pi}$ is $2 - \beta - \gamma$. We establish a slightly more general inequality in Theorem 2 below. To make the connection, imagine choosing $i = 1, \ldots, n$ at random. The random variable ζ is q_i and η is s_i. We have defined $\pi_i = 0$ when $q_i + s_i = 0$; by a separate little argument, we may restrict $q_i + s_i$ to be positive. To get the upper bound on $\overline{\pi}$, take $\overline{q} = \beta$ and $\overline{s} = \beta + \gamma - 1$; then a in Theorem 2 is $\overline{q} = 1 - \beta$ and $b = \overline{s} = \beta + \gamma - 1$.

Theorem 2. Fix non-negative real numbers a, b with $a + b < 1$. Suppose ζ, η are random variables with $0 \leq \zeta, \eta \leq 1$, $0 < \zeta + \eta \leq 1$, $E\{\zeta\} = a$, and $E\{\eta\} = b$. Then $E\{\zeta/(\zeta + \eta)\} \leq 1 - b$, and the bound is sharp.

Proof. We consider the upper bound first. Without real loss of generality, suppose $\zeta > 0$. Construct $\eta' = 0$ or $1 - \zeta$ with $E\{\eta'|\zeta\} = E\{\eta|\zeta\}$. Then $E\{\eta'\} = E\{\eta\} = b$, and $E\{\zeta/(\zeta + \eta)\} \leq E\{\zeta/(\zeta + \eta')\}$, because the function $y \to x/(x + y)$ is convex on $[0, \infty)$ for each $x > 0$. Thus, it suffices to consider ζ, η such that $\eta = 0$ or $1 - \zeta$. If $\eta = 0$, then $\zeta/(\zeta + \eta) = 1$; on the other hand, if $\eta > 0$, then $\eta = 1 - \zeta$ and $\zeta/(\zeta + \eta) = \zeta = 1 - \eta$. Now

$$\begin{aligned} E\{\zeta/(\zeta + \eta)\} &= P\{\eta = 0\} + E\{\zeta|\eta > 0\}P\{\eta > 0\} \\ &= P\{\eta = 0\} + E\{1 - \eta|\eta > 0\}P\{\eta > 0\} \\ &= P\{\eta = 0\} + P\{\eta > 0\} - E\{\eta|\eta > 0\}P\{\eta > 0\} \\ &= 1 - E\{\eta\} \\ &= 1 - b, \end{aligned}$$

which completes the proof of the upper bound. Any pair of non-negative random variables ζ, η with the given expectations and $\eta = 0$ or $1 - \zeta$ will achieve the bound, as the display shows. For a specific example, let $\zeta = a$ and let η be 0 or $1 - a$ with probabilities $(1-a-b)/(1-a)$ and $b/(1-a)$ respectively. Then $E\{\zeta\} = a$, $E\{\eta\} = b$, and $E\{\zeta/(\zeta+\eta)\} = (1 - a - b)/(1 - a) + ab/(1 - a) = 1 - b$, as required.

PROBABILITY OF SPECIFIC CAUSATION

We now address the question, "what is the chance that exposure caused injury?" Since the crucial parameters are not identifiable, this question does not have an answer. It may be suggested in response that, absent information to the contrary, the plaintiff should be viewed as randomly picked from the study population. However, random selection can be operationalized in a variety of ways, and the probability of causation depends quite strongly

on the details of the selection mechanism. That is what we show next, by considering three scenarios for choosing a subject from the study population.

Scenario 1. Pick a subject at random; condition that the subject is among the exposed and injured.

Scenario 2. Divide the subjects at random into two groups, exposed and unexposed. Condition that at least one of the exposed subjects is injured. Pick a subject at random from those who are exposed and injured.

Scenario 3. Pick a subject at random; condition that the subject is among the exposed and injured; furthermore, condition that the subject sues.

As indicated earlier, the three scenarios share maintained hypotheses:

(7) m out of the n subjects are assigned at random to exposure; the assignment variables $\{X_i\}$ are independent of the response variables $\{U_i, V_i\}$; the response variables are independent across subjects i.

SCENARIO 1

Let ξ be a random integer between 1 and n, indexing our randomly selected subject. In scenario 1, we are conditioning that $X_\xi = V_\xi = 1$, that is, subject ξ is among the exposed and injured. Recall that $\mathrm{RR} = \gamma/\beta$ is the relative risk.

Proposition 1. Suppose (1)–(2)–(3) and (7). Then

$$P\{U_\xi = 0 | X_\xi = 1, V_\xi = 1\} \geq 1 - \frac{1}{\mathrm{RR}}.$$

Proof. By assumption (7), X_k is independent of (U_k, V_k) and $P\{X_k = 1\} = m/n$. Now

$$\begin{aligned} P\{U_\xi = 0, X_\xi = 1, V_\xi = 1\} &= \sum_{k=1}^{n} P\{U_k = 0, X_k = 1, V_k = 1, \xi = k\} \\ &= \frac{1}{n} \sum_{k=1}^{n} P\{U_k = 0, X_k = 1, V_k = 1\} \\ &= \frac{m}{n} \frac{1}{n} \sum_{k=1}^{n} P\{U_k = 0, V_k = 1\} \\ &= \frac{m}{n}\, \overline{q}. \end{aligned} \tag{8}$$

Thus $P\{U_\xi = 0, X_\xi = 1, V_\xi = 1\} \geq m(\gamma - \beta)/n$ by Lemma 1b. Similarly,

$$P\{X_\xi = 1, V_\xi = 1\} = m\gamma/n. \tag{9}$$

Finally, $\dfrac{\gamma - \beta}{\gamma} = 1 - \dfrac{1}{\mathrm{RR}}$, which completes the proof.

The right side of the inequality in Proposition 1 is called the "etiologic fraction," or "attributable risk." According to Proposition 1, if the relative risk is greater than 2.0, the chance that a randomly selected subject was injured by exposure—given that the subject was exposed and injured—exceeds 50%. That seems to be the best argument connecting relative risk with the probability of causation. In scenario 1, the upper bound on the probability of causation is easily seen to be 1, by Lemma 1(a). Proposition 1 appears in Robins and Greenland (1989ab), with extensions to time-dependent risks. Also see Beyea and Greenland (1999) and Pearl (1999). Other bounds are discussed by Manski (1995, chapter 4).

SCENARIO 2

Let $\mathcal{R}$ denote the subjects who are exposed and injured. Thus, $\mathcal{R} \subset \mathcal{X}$, the latter being the set of subjects who are exposed. By assumption, $\mathcal{X}$ is random and $|\mathcal{X}| = m$, where $|J|$ is the number of elements in J. Let ρ be uniformly distributed over $\mathcal{R}$ when $\mathcal{R}$ is nonempty. Thus, $U_\rho = 0$ is the event that exposure caused injury, for a person ρ randomly selected from the subjects who are exposed and injured. For the next result, let J be a typical value of nonempty $\mathcal{R}$, that is, a subset of $\{1, \ldots, n\}$ with $1 \le |J| \le m$. The calculation looks somewhat technical, but the point is simple: scenarios 1 and 2 seem very similar at first reading, but lead to different probabilities of causation.

Proposition 2. Suppose (1)–(2)–(3) and (7); the π_i are defined by (5). Then

$$P\{U_\rho = 0 | \mathcal{R} \neq \emptyset\} = \Big[\sum_J \Big(\frac{1}{|J|}\sum_{j\in J}\pi_j\Big)P\{\mathcal{R} = J\}\Big]\Big/\Big[\sum_J P\{\mathcal{R} = J\}\Big].$$

Proof. The key step is

$$\begin{aligned} P\{U_\rho = 0, \mathcal{R} = J\} &= \sum_{j\in J} P\{\rho = j, U_j = 0, \mathcal{R} = J\} \\ &= \frac{1}{|J|}\sum_{j\in J} P\{U_j = 0, \mathcal{R} = J\} \\ &= \frac{1}{|J|}\sum_{j\in J} \pi_j P\{\mathcal{R} = J\}; \end{aligned} \tag{10}$$

the last equality holds because $j \in \mathcal{R} = J$ entails $X_j = V_j = 1$, and

$$P\{U_j = 0 | \mathcal{R} = J\} = P\{U_j = 0 | X_j = V_j = 1\} = P\{U_j = 0 | V_j = 1\} = \pi_j$$

by (5) and (7). The balance of the proof is omitted as routine.

To compare scenarios 1 and 2 in a simple case, suppose $n = 3$ and $m = 2$. As before, let $\gamma_i = q_i + s_i$, the probability of injury if exposed. The conditional probability of "no injury if unexposed" given "exposed and injured" for subject ξ in scenario 1 is

$(q_1+q_2+q_3)/(\gamma_1+\gamma_2+\gamma_3)$, by (8) and (9). This is a weighted average of π_1, π_2, π_3, the weights being

$$\gamma_1,\ \gamma_2,\ \gamma_3.$$

The conditional probability of "no injury if unexposed" given "exposed and injured" for subject ρ in scenario 2 is also a weighted average of π_1, π_2, π_3 but with different weights:

$$\gamma_1\Big(2-\frac{3}{2}\gamma+\frac{1}{2}\gamma_1\Big),\ \gamma_2\Big(2-\frac{3}{2}\gamma+\frac{1}{2}\gamma_2\Big),\ \gamma_3\Big(2-\frac{3}{2}\gamma+\frac{1}{2}\gamma_3\Big).$$

As before, $\gamma=(\gamma_1+\gamma_2+\gamma_3)/3$. Of course, the random index is differently defined in the two scenarios.

The result for scenario 2 follows from (10), because

$$P\{U_\rho=0\}=\sum_{j=1}^n \omega_j\pi_j$$

where

$$\omega_j=\sum_{J:j\in J}\frac{1}{|J|}P\{\mathcal{R}=J\}. \tag{11}$$

If $j=1$, say, then $J=\{1\}$ or $\{1,2\}$ or $\{1,3\}$. Now $\mathcal{R}=\{1\}$ iff either

$$\mathcal{X}=\{1,2\} \text{ and } V_1=1, V_2=0$$

or

$$\mathcal{X}=\{1,3\} \text{ and } V_1=1, V_3=0.$$

So $P\big\{\mathcal{R}=\{1\}\big\}=[\gamma_1(1-\gamma_2)+\gamma_1(1-\gamma_3)]/3$. Similarly, $P\big\{\mathcal{R}=\{1,2\}\big\}=\gamma_1\gamma_2/3$ and $P\big\{\mathcal{R}=\{1,3\}\big\}=\gamma_1\gamma_3/3$. By (11),

$$\omega_1=\frac{\gamma_1}{3}\Big((1-\gamma_2)+(1-\gamma_3)+\frac{1}{2}\gamma_2+\frac{1}{2}\gamma_3\Big)=\frac{\gamma_1}{6}\Big(4-\gamma_2-\gamma_3\Big)=\frac{\gamma_1}{3}\Big(2-\frac{3}{2}\gamma+\frac{1}{2}\gamma_1\Big),$$

as required.

If n is large and the probabilities not ill-behaved, it must be noted that the difference between the results in the two scenarios will be small. Indeed, the right hand side of Proposition 2 can be rewritten as

$$\sum_{i=1}^n \pi_i\gamma_i w_i\Big/\sum_{i=1}^n \gamma_i w_i$$

where $w_i=E\{1/(1+N_i)\}$, the random variable N_i being the number of injured persons among the exposed if subject i is removed from the study population, and $m-1$ of the remaining $n-1$ persons are assigned at random to exposure; to make the connection with (11), as will be seen below, $\omega_i=m\gamma_i w_i/n$. It will also be seen that the

w_i are essentially all the same, so the weighted average for scenario 2 almost reduces to $\sum_i \pi_i\gamma_i / \sum_i \gamma_i = \sum_i q_i / \sum_i \gamma_i$, just as for scenario 1.

In more detail, fix i. Let $\mathcal{R}_i$ be the random set of exposed and injured persons in the smaller study with subject i removed. So $N_i = |\mathcal{R}_i|$. Let J_i run over the subsets of $\{1, \ldots, n\} - \{i\}$ with $0 \le |J_i| \le m-1$. As before, $i \in J$ and $\mathcal{R} = J$ entails $X_i = V_i = 1$, so $P\{\mathcal{R} = J\} = (m\gamma_i/n)P\{\mathcal{R} = J | X_i = V_i = 1\}$. Then

$$\begin{aligned}\sum_{J:i\in J} \frac{1}{|J|} P\{\mathcal{R} = J\} &= \frac{m}{n}\gamma_i \sum_{J:i\in J} \frac{1}{|J|} P\{\mathcal{R} = J | X_i = V_i = 1\} \\ &= \frac{m}{n}\gamma_i \sum_{J_i} \frac{1}{1+|J_i|} P\{\mathcal{R}_i = J_i\} \\ &= \frac{m}{n}\gamma_i E\Big\{\frac{1}{1+N_i}\Big\};\end{aligned}$$

in the first line, the sum runs over sets J with $i \in J$; in the second, over J_i. Now, for instance, the numerator on the right hand side of the equality in Proposition 2 can be written as

$$\sum_{i=1}^n \pi_i \sum_{J:i\in J} \frac{1}{|J|} P\{\mathcal{R} = J\} = \frac{m}{n} \sum_{i=1}^n \pi_i\gamma_i w_i,$$

as required: m/n will cancel in numerator and denominator.

Let N be the number of injured persons among the exposed in the original design, with m at random out of n assigned to treatment. Make the additional assumption that

$$\lambda = E\{N\} = m\gamma \text{ is large;}$$

if not, the study would have little statistical power. Write $x \approx y$ if $x/y \to 1$, and $x = O(y)$ if x/y stays bounded.

Lemma 2. In (b-c-d), let $\lambda \to \infty$.
(a) $\operatorname{var} N \le \lambda$.
(b) $P\{|N - \lambda| > \delta\lambda\} = O(\rho^\lambda)$, where $0 < \rho < 1$ depends on $\delta > 0$.
(c) $\int 1/(1+N)\, dP \approx 1/\lambda$.
(d) $\int_{N\ge 1} 1/[N(1+N)]\, dP \approx 1/\lambda^2$.

Proof. Claim (a). By straightforward but tedious calculation,

$$\operatorname{var} N = m\gamma(1-\gamma) - \frac{m(m-1)}{n-1}\sigma^2, \quad \text{with } \sigma^2 = \frac{1}{n}\sum_{i=1}^n (\gamma_i - \gamma)^2.$$

Claim (b) follows from results in Hoeffding (1963); also see Freedman (1973). Explicit bounds are possible. Claims (c) and (d) are immediate from (b).

Recall that N is the number of injured persons among the exposed in the original design, with m at random out of n assigned to treatment; on the other hand, N_i is the

number of injured persons among the exposed if subject i is removed from the study population, and $m-1$ of the remaining $n-1$ persons are assigned at random to exposure. The next lemma shows that results for scenarios 1 and 2 are close, provided λ is large.

Lemma 3. Suppose λ tends to infinity. Then $w_i \approx 1/\lambda$ uniformly in i.

Proof. We can couple N_i and N so that $N - N_i = 0$ or 1. Indeed, suppose $n \geq m \geq 1$. Construct N in the usual way: choose m subjects out of n for the exposure group $\mathcal{X}$, and let $N = \sum_{j \in \mathcal{X}} V_j$. To construct N_i, let $\zeta = i$ if $i \in \mathcal{X}$; otherwise, let ζ be uniformly distributed over $\mathcal{X}$. Let

$$\mathcal{X}_i = \mathcal{X} - \{\zeta\} \text{ and } N_i = \sum_{j \in \mathcal{X}_i} V_j = N - V_\zeta.$$

By symmetry, $\mathcal{X}_i$ is a random sample of size $m-1$ from $\{1, \ldots, n\} - \{i\}$, which completes the coupling.

With this coupling, $N = 0$ entails $N_i = 0$. Furthermore,

$$\frac{1}{1+N} \leq \frac{1}{1+N_i} \leq \frac{1}{1+N} + \frac{Z}{N(1+N)},$$

where $Z = 1$ if $N \geq 1$, $Z = 0$ if $N = 0$, and $0/0 = 0$. Lemma 2 completes the proof.

SCENARIO 3

Suppose that an exposed, injured person files suit with probability depending on individual characteristics, with healthier persons more likely to sue. More specifically, there is an additional random variable Y_i for each subject. If $X_i = 1$ and $V_i = 1$, so that i was exposed and injured, then $Y_i = 0$ or 1, and $\{Y_i = 1\}$ corresponds to the event that i files suit. There is no need to define Y_i if $X_i = 0$ or $V_i = 0$. We assume

$$P\{Y_i = 1 | X_i = V_i = 1\} = \frac{\lambda}{q_i + s_i}, \tag{12}$$

where λ is a positive constant. The following result is easily proved, following the pattern of (8).

Proposition 3. Suppose $q_i + s_i > 0$ for all i; suppose (1)–(2)–(3) and (7); the π_i are defined by (5); the probability of suit given injury and exposure is (12). Let ξ be a random integer between 1 and n. Then

$$P\{U_\xi = 0 | X_\xi = V_\xi = Y_\xi = 1\} = \overline{\pi}.$$

Proposition 3 brings the average probability of causation back into play. Individual differences in the propensity to sue and the chance of injury can make the probability of specific causation arbitrarily small, even for arbitrarily large values of the relative risk.

That is the consequence of Proposition 3 and Theorem 1. Scenarios 1 and 3 give very different answers for the probability of causation, illustrating yet again one of the basic weaknesses in the principle of insufficient reason. (The principle of insufficient reason holds that ignorance should be expressed by uniformity, and one standard objection is that uniformity can usually be expressed in many different ways; other—perhaps more basic—objections will not be discussed here.)

DISCUSSION

The model seems useful in that it provides a clear definition for the probability of causation π_i, and formalizes the role of individual differences: the response probabilities p_i, q_i, r_i, s_i depend on the individual i. Limitations of relative risk in assessing evidence of specific causation then become apparent. (i) Epidemiologic data cannot restrict the probability of causation to any useful degree: the key parameters are not identifiable, and even the average probability of causation is subject to wide uncertainties, so estimates generally depend on assumptions that are not testable. (ii) Even if we grant that plaintiff in a law case should be treated as a randomly chosen member of a study population, results depend strongly on the details: exactly how is random choice to be implemented?

Of course, reality is substantially more complicated than the model. Assignment to exposure is likely to correlate with probable responses (the problem of confounding), reported outcomes are likely to differ from true outcomes, and so forth. Moreover, the probability of a counterfactual hypothetical does not fit with any ease into the classical frequentist architecture of statistics. The complexities are unlikely to strengthen the case for relative risk in assessing specific causation.

APPENDIX 2

BACKGROUND RATES AND TABLE 1

Langmuir et al. (1984) classify cases as types A–E: types A–D refer to extent of paralysis, from most extensive to least; type E had insufficient data. "Extensive" cases in the first row of Table 1 are types A and B. Following Langmuir et al., background rates are truncated below at 0.14 for extensive cases and 0.21 for cases of types A–D; from data in the paper, we add 0.03 for the background rate of type E cases, giving a total background rate of 0.24 per million person-weeks of exposure; compare Tr. 17.44. In the second column of Table 1, we get 7.40 for the expected number of GBS cases of all types, compared to the 7.45 reported in note 10 of the opinion. The procedure for computing expected numbers is discussed next.

COMPUTING THE RELATIVE RISK OF GBS

Observation periods are indexed by $k = 1, \ldots, 17$. To get started, assume each period is a week long. The following data are recorded in Langmuir et al. (1984):

(a) N_k, the number of persons vaccinated in period k;
(b) m_k, the number of new cases of GBS among unvaccinated persons in period k;

(c) o_ℓ, the number of new cases of GBS among vaccinated persons in the ℓth week after vaccination.

The GBS cases are broken down by type. Given the total population at the beginning of the vaccination campaign, Langmuir et al. estimate M_k, the number of unvaccinated persons in period k. (Attention is confined to persons 18 years or older.)

We assume that subjects vaccinated in the kth period were vaccinated in the middle of that period; unvaccinated subjects who contracted GBS in the kth period contracted GBS in the middle of that period; vaccinated subjects who contracted GBS in the ℓth week after vaccination did so in the middle of that week. For example, subjects vaccinated in period k who contracted GBS in the ℓth week after vaccination are considered to have been vaccinated at week $k - 0.5$ and to have contracted GBS $\ell - 0.5$ weeks later, at the end of the $k + \ell - 1$st period.

Suppose vaccination did not change the risk of GBS. On this null hypothesis, the expected number of new GBS cases among the vaccinated in the ℓth week after vaccination can be computed as a sum, with contributions from the number vaccinated in each period, times the background rate among the unvaccinated ℓ weeks later. We estimate the background rate at the end of the kth period by averaging rates in the kth and $k + 1$st periods:

$$b_k = \frac{1}{2}\Big(\frac{m_k}{M_k} + \frac{m_{k+1}}{M_{k+1}}\Big). \tag{13}$$

Let e_ℓ be the expected number of new GBS cases among vaccinated persons in the ℓth week after vaccination, computed as follows:

$$e_\ell = \sum_{k=1}^{17-\ell} N_k b_{k+\ell}. \tag{14}$$

There are no vaccinations after the moratorium in the 11th period, so $N_{12} = \cdots = N_{17} = 0$. And $\ell = 1, \ldots, 16$: one period gets used up in averaging the background rates. Equation (14) stratifies on time of vaccination: there is one term in the sum for each vaccination cohort, and the expected number of GBS cases among the vaccinated in the ℓth week after vaccination is the sum of the contributions from each cohort. The relative risk in the ℓth period after vaccination is

$$r_\ell = o_\ell / e_\ell. \tag{15}$$

Generally, the periods are weeks; period 1 is 10 days long and period 17 is 8 days. Langmuir et al. adjust for longer periods by converting counts to "person-weeks" at risk. Persons vaccinated in period 1 should be treated as having been vaccinated at day 5, so our use of half-periods is slightly inconsistent; the axis labels in Figure 1—for background rates—must be charitably interpreted. In Table 1, the expected numbers are $e_{11} + \cdots + e_{16}$. Background rates are truncated below in column 1 or left alone (column 2). And in the last line of Table 1, 8 cases are added to $o_{11} + \cdots + o_{16}$ to implement the effect of sanctions.

The ratio in (15) is more properly called a "rate ratio," as it is the ratio of two incidence rates.

REPORTING OF VACCINATED GBS CASES

To examine under- or over-reporting of vaccinated GBS cases, we smoothed the relative risk curve in Figure 2b by rounding in periods 1–6; we then allowed the curve to decay linearly at the rate of 0.1 per period, so the relative risk becomes 1.0 at period 17. In periods 7–14, the smoothed curve is generally higher than the real one; in periods 15–16, there a few late-onset cases which bring up the real curve. Given the smooth curve, which is based largely on pre-moratorium data, we can project the number of GBS cases among the vaccinated by calendar period rather than week since vaccination. In period 10, we project 61 cases and observe 71; in period 11, we project 51 and observe 61. On the other hand, in periods 12–17, we project 138 and observe 111. (Observations are in Table 5 of Langmuir et al.) The pattern is suggestive, but its statistical significance seems marginal.

Our projections are computed on the following basis. Let s_ℓ be the smoothed relative risk for GBS in the ℓth week after vaccination; let b_k be the background rate at the end of period k, as in (13); let

$$p_k = \sum_{j=1}^{k-1} N_j s_{k-j} b_{k-1}. \tag{16}$$

Then p_k is the projected number of GBS cases among the vaccinated, with onsets at the end of calendar period $j - 0.5 + k - j - 0.5 = k - 1$. The projected number of GBS cases with onsets in calendar period k is

$$\frac{1}{2}(p_k + p_{k+1}). \tag{17}$$

THE EFFECT OF PRIOR ILLNESS

We proceed informally. Let I be the event of prior illness, V the event of vaccination, and G the event of GBS. Write F^c for the complement of F. Let

$$P(I|GV) = \alpha_1, \quad P(I|GV^c) = \alpha_2, \quad P(G|V) = \lambda P(G|V^c).$$

In this formalism, λ is the analog of the relative risk. Goldfield and Mantel seemed to be making the following argument:

$$P(G|IV) = \frac{P(IGV)}{P(IV)} = \frac{\alpha_1 P(GV)}{P(IV)},$$
$$P(G|IV^c) = \frac{P(IGV^c)}{P(IV^c)} = \frac{\alpha_2 P(GV^c)}{P(IV^c)}.$$

Therefore,

$$\frac{P(G|IV)}{P(G|IV^c)} = \frac{\alpha_1}{\alpha_2}\frac{P(GV)}{P(IV)}\frac{P(IV^c)}{P(GV^c)} = \frac{\alpha_1}{\alpha_2}\frac{P(G|V)}{P(I|V)}\frac{P(I|V^c)}{P(G|V^c)} = \frac{\alpha_1}{\alpha_2}\lambda\frac{P(I|V^c)}{P(I|V)}.$$

That is why the relative risk should be adjusted by the factor α_1/α_2.

In our first application of this argument, $\alpha_1 = .33$ and $\alpha_2 = .62$; there is a tacit assumption that $P(I|V^c) \doteq P(I|V)$. The objection raised by Goldfield and Mantel is that $P(I|V^c) > P(I|V)$: vaccination is contra-indicated right after illness. Of course, the meaning of I has shifted slightly during the argument, from illness just prior to the onset of GBS to illness just prior to vaccination, or illness in general. Moreover, if persons susceptible to illness are more likely to seek vaccination—as seems plausible—the selection effect may go the other way.

ACKNOWLEDGMENTS

We would like to thank the following persons for useful discussions: Michael Berger (NIH), Richard Berk (UCLA), Joe Cecil (Washington, D.C.), Mike Finkelstein (New York), Mike Green (University of Iowa), Angelika Hahn (University of Western Ontario), Paul Humphreys (University of Virginia), James Robins (Harvard), Larry Schonberger (CDC). Many of the participants in the case generously shared their knowledge with us, including three of the epidemiology experts (Leonard Kurland, Nathan Mantel, Neal Nathanson) and the two lawyers who argued the epidemiologic evidence (Leslie Ohta, Charles Thomas). The Department of Justice kindly provided surviving portions of the trial transcript. Part of this work was completed while PBS was on appointment as a Miller Research Professor in the Miller Institute for Basic Research in Science.

REFERENCES

American Medical Association, Council on Scientific Affairs. 1987. Radioepidemiological tables. *Journal of the American Medical Association* 257: 806–09.

Bailey, L. A., L. Gordis, and M. Green. 1994. Reference guide on epidemiology. *Reference Manual on Scientific Evidence*. Federal Judicial Center.

Beyea, J. and S. Greenland. 1999. The importance of specifying the underlying biologic model in estimating the probability of causation. *Health Physics* 76: 269–74.

Beghi, E., L. T. Kurland, D. W. Mulder, and W. C. Wiederholt. 1985. Guillain-Barré syndrome: clinicoepidemiologic features and effect of influenza vaccine. *Archives of Neurology* 42: 1053–7.

Black, B., and D. E. Lilienfeld. 1984. Epidemiologic proof in toxic tort litigation. *Fordham Law Review* 52: 732–85.

Breman, J. G., and N. S. Hayner. 1984. Guillain-Barré syndrome and its relationship to swine influenza vaccination in Michigan, 1976–1977. *American Journal of Epidemiology* 119: 880–9.

Freedman, D. 1973. Another note on the Borel-Cantelli lemma and the strong law with the Poisson approximation as a byproduct. *Annals of Probability* 1: 910–25.

Freedman, D. 1999. From association to causation: some remarks on the history of statistics. Technical report 537, Department of Statistics, UC Berkeley. To appear in *Statistical Science*.

Hahn, A. F. 1998. Guillain-Barré syndrome. *Lancet* 352: 635–41.

Hart, H. L. A. and A. M. Honoré. 1985. *Causation in the Law*. 2nd edition. Oxford.

Hoeffding, W. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58: 13–30.

Hodges, J. L. Jr., and E. Lehmann. 1964. *Basic Concepts of Probability and Statistics*. Holden-Day, San Francisco (section 9.4).

Holland, P. 1988. Causal inference, path analysis, and recursive structural equations models. In *Sociological Methodology*, edited by C. Clogg. American Sociological Association, Washington, D. C., Chapter 13.

Hughes, R. A. 1990. *Guillain-Barré Syndrome*. Springer-Verlag, London.

Hughes, R. A., and J. H. Rees. 1997. Clinical and epidemiologic features of Guillain-Barré syndrome. *Journal of Infectious Diseases* 176 Suppl 2: S92–8.

Hurwitz, E. S., L. B. Schonberger, D. B. Nelson, and R. C. Holman. 1981. Guillain-Barré syndrome and the 1978–1979 influenza vaccine. *New England Journal of Medicine* 304: 1557–61.

Johnson, M. T., C. E. Drew, and D. P. Miletich. 1998. *Use of Expert Testimony, Specialized Decision Makers, and Case-Management Innovations in the National Vaccine Injury Compensation Program.* Federal Judicial Center, Washington, D. C.

Kaplan, J. E., P. Katona, E. S. Hurwitz, and L. B. Schonberger. 1982. Guillain-Barré syndrome in the United States, 1979–1980 and 1980–1981: lack of an association with influenza vaccination. *Journal of the American Medical Association* 248: 698–700.

Kurland, L. T., W. C. Wiederholt, J. W. Kirkpatrick, H. G. Potter, and P. Armstrong. 1985. Swine influenza vaccine and Guillain-Barré syndrome. Epidemic or artifact? *Archives of Neurology* 42: 1089–90.

Langmuir, A. D. 1979. Guillain-Barré syndrome: the swine influenza virus vaccine incident in the United States of America, 1976–77: preliminary communication. *Journal of the Royal Society of Medicine* 72: 660–9.

Langmuir, A. D., D. J. Bregman, L. T. Kurland, N. Nathanson, and M. Victor. 1984. An epidemiologic and clinical evaluation of Guillain-Barré syndrome reported in association with the administration of swine influenza vaccines. *American Journal of Epidemiology* 119: 841–79.

Lasky, T., G. J. Terracciano, L. Magder, C. L. Koski, M. Ballesteros, D. Nash, S. Clark, P. Haber, P. D. Stolley, L. B. Schonberger, and R. T. Chen. 1998. The Guillain-Barré syndrome and the 1992–1993 and 1993–1994 influenza vaccines. *New England Journal of Medicine* 339: 1797–1802.

Manski, C. F. 1995. *Identification Problems in the Social Sciences*. Harvard University Press, Cambridge, Mass.

Manko v. United States. 636 F.Supp. 1419 (W.D.Mo. 1986). Aff'd in part, 830 F.2d 831 (8th Cir. 1987).

Mantel, N. 1985. Re: "An epidemiologic and clinical evaluation of Guillain-Barré syndrome reported in association with the administration of swine influenza vaccines" [letter]. *American Journal of Epidemiology* 121: 620–3. With a reply by Langmuir et al., 621–23.

Marks, J. S., and T. J. Halpin. 1980. Guillain-Barré syndrome in recipients of A/New Jersey influenza vaccine. *Journal of the American Medical Association* 243: 2490–4.

Nachamkin, I., B. M. Allos, and T. Ho. 1998. Campylobacter species and Guillain-Barré syndrome. *Clinical Microbiology Reviews* 11: 555–67.

Neustadt, R. E., and H. V. Fineberg. 1982. *The Epidemic That Never Was: Policy-Making and the Swine Flu Affair*. Random House, New York.

Neyman, J. 1923. Sur les applications de la théorie des probabilités aux experiences agricoles: Essai des principes. *Roczniki Nauk Rolniczki* 10: 1–51, in Polish; English translation by Dabrowska, D. and T. Speed. 1990. *Statistical Science* 5: 463–80.

Pearl, J. 1999. Probabilities of causation: three counterfactual interpretations and their identification. Technical report, Department of Computer Science, UCLA. To appear in *Synthèse*.

Petitti, D. B. 1996. Review of "Reference Guide on Epidemiology." *Jurimetrics* 36: 159–68.

Robins, J., and S. Greenland. 1989a. Estimability and estimation of excess and etiologic fractions. *Statistics in Medicine* 8: 845–59.

Robins, J., and S. Greenland. 1989b. The probability of causation under a stochastic model for individual risk. *Biometrics* 45: 1125–38.

Ropper, A. H., E. F. M. Wijdicks, and B. T. Truax. 1991. *Guillain-Barré Syndrome*. F. A. Davis, Philadelphia.

Rubin, D. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66: 688–701.

Safranek, T. J., D. N. Lawrence, L. T. Kurland, D. H. Culver, W. C. Wiederholt, N. S. Hayner, M. T. Osterholm, P. O'Brien, J. M. Hughes, and the Expert Neurology Group. 1991. Reassessment of the association between Guillain-Barré syndrome and receipt of swine influenza vaccine in 1976–1977: results of a two-state study. *American Journal of Epidemiology* 133: 940–51.

Schonberger, L. B., D. J. Bregman, J. Z. Sullivan-Bolyai, R. A. Keenlyside, D. W. Ziegler, H. F. Retailliau, D. L. Eddins, and J. A. Bryan. 1979. Guillain-Barré syndrome following vaccination in the National Influenza Immunization Program, United States, 1976–1977. *American Journal of Epidemiology* 110: 105–23.

Silverstein, A. M. 1981. *Pure Politics and Impure Science: The Swine Flu Affair*. Johns Hopkins University Press, Baltimore.