

The use of the ‘reverse Cornfield inequality’ to assess the sensitivity of a non-significant association to an omitted variable

Binbing Yu^{1,*},[†] and Joseph L. Gastwirth^{2,3}

¹*Information Management Services, Inc., 12501 Prosperity Dr. Suite 200, Silver Spring, MD 20904, U.S.A.*

²*Division of Cancer, Epidemiology and Genetics, EPS 8804, National Cancer Institute, Bethesda, MD 20892, U.S.A.*

³*Department of Statistics, The George Washington University, 2201 G St. NW, Washington DC 20052, U.S.A.*

SUMMARY

Unlike randomized experimental studies, investigators do not have control over the treatment assignment in observational studies. Hence, the treated and control (non-treated) groups may have widely different distributions of unobserved covariates. Thus, if observational data are analysed as if they had arisen from a controlled study, the analyses are subject to potential bias. Sensitivity analysis is a technique for assessing whether the inference drawn from a study could be altered by a moderate ‘imbalance’, between the distribution of the covariates in different groups. In this paper, we examine the sensitivity analysis of the test of proportions in 2×2 tables from a new perspective: ‘could a non-significant result have occurred because the treated group has a higher prevalence of an unobserved risk factor?’. The study was motivated by an analysis of the studies concerning with the possible effect of spermicide use on birth defects that were cited in a legal decision. Copyright © 2003 John Wiley & Sons, Ltd.

1. INTRODUCTION

In his description of the late Prof. Cornfield’s contributions, Greenhouse [1] emphasized the importance of his result assessing the potential impact of an omitted variable (OV) on a positive finding. Cornfield’s inequality [2] gave conditions that an OV, U say, had to satisfy in order that an observed relative risk, R , of the agent under study could be entirely due to different prevalences of U in the two groups, i.e. the relative risk of U must exceed R and the ratio of the prevalence of U in the exposed group to that of the unexposed group must exceed R . Originally, the inequality was used to show that the strong association between tobacco use and lung cancer was very unlikely to be explained by another factor. Later Gastwirth and

* Correspondence to: Binbing Yu, Information Management Services, Inc., 12501 Prosperity Dr. Suite 200, Silver Spring, MD 20904, U.S.A.

[†] E-mail: yub@ims.nci.nih.gov

Contract/grant sponsor: NSF

Copyright © 2003 John Wiley & Sons, Ltd.

Received May 2003

Accepted May 2003

Greenhouse [3] showed that the methodology was useful in the analysis of data arising in employment discrimination cases.

Today, epidemiologic studies are admitted as evidence in cases concerning possible harm caused by drugs or chemical exposures. Often the original inequality is used to assess whether another agent could have created an association that is shown in studies indicating that the chemical at issue caused the harm. In our case, the defendant relied on a study that showed a non-significant increased risk. There were other known or suggested risk factors that were more prevalent in the controls. This example motivated us to create a 'reverse' inequality that enabled us to examine whether an unobserved covariate might increase a non-significant risk factor to 'significance'. Li *et al.* [4] used a sensitivity analysis for a similar purpose. In the context of equivalence testing, they rejected the hypothesis that treatment was very much better than control and found this absence of effect to be insensitive to hidden bias.

In Section 2 we review Cornfield's original inequality and some useful extensions. The reverse inequality, which is appropriate when the prevalence of the omitted risk factor is higher in the control group so that the observed relative risk underestimates the true risk, is presented in Section 3. Results for effects that follow an additive or multiplicative or logistic model are given and illustrated. Section 4 describes the legal case and study that motivated our research. Here, the reverse inequality is applied to the case concerning the potential effect of spermicide use on limb reduction defects. In the study, a higher fraction of control mothers were likely to have been exposed to a previously known risk factor for which data were not collected. As the observed relative risk was greater than one, we were concerned that the imbalance between the case and control groups with respect to other risk factors might have reduced a statistically significant risk to a non-significant one. Thus, the purpose of our analysis can be regarded as the 'reverse' of the original one.

2. CORNFIELD'S INEQUALITY

In response to claims that the relationship between smoking and lung cancer could be explained by a genetic or other omitted variable, Cornfield *et al.* [2] developed an inequality relating the observed relative risk to the imbalance in prevalence of the omitted variable in smoking and non-smoking groups and the relative risk of the omitted variable.

Formally, Cornfield's inequality involves three binary variables: (1) $Y = 1$ (0) for positive (negative) response, (2) $X = 1$ (0) for exposed (unexposed) group, (3) $U = 1$ (0) for presence (absence) of the unobserved variable. Assume that the prevalences of U are $f_1 = P(U = 1 | X = 1)$ in the exposed group and $f_0 = P(U = 1 | X = 0)$ in the control group, respectively. Let

$$\pi_{x,u} = P(Y = 1 | X = x, U = u)$$

$$\pi_x = P(Y = 1 | X = x)$$

be the response probabilities, where x and u equal 0 or 1. If U is omitted, the observed data are summarized in a single 2×2 table and we only observe π_0 and π_1 with relative risk $R^* = \pi_1/\pi_0$. Here, π_1 and π_0 are the probabilities of positive response in the exposed and control groups, respectively.

Table I. Limb defects and spermicide usage after last menstrual period.

Limb defect	Spermicide usage	
	Non-user (0)	User (1)
Limb defect (1)	$a = 75$	$b = 75$
No limb defect (0)	$c = 2756$	$d = 2208$
Total	$n_0 = 2831$	$n_1 = 2283$

Source: Table III in Mills *et al.* [5].

Mills *et al.* [5] examined the relationship between maternal spermicide use and congenital malformations. Table I summarizes the number of limb defect(Y) by spermicide use(X) from the study of Mills *et al.* In Table I, the estimated birth defect fractions are $\hat{\pi}_0 = \frac{75}{2831} = 0.0265$ and $\hat{\pi}_1 = \frac{75}{2283} = 0.0329$. The observed relative risk $R^* = \pi_1/\pi_0$ is 1.24. The original analysis included some additional factors that did not change the estimated relative risk of spermicide use.

Could an unobserved variable U fully explain the observed relative risk, R^* , which appears to be due to agent X ? If this were the case, then Y would be independent of X given U , i.e. $P(Y = 1|X = 1, U = u) = P(Y = 1|X = 0, U = u) = P(Y = 1|U = u)$ for both $u = 0$ and 1. Hence the relative risk, R_U , associated with U is equal to $P(Y = 1|U = 1)/P(Y = 1|U = 0)$. Cornfield showed that for this to occur, the following two conditions must hold:

$$R_U \geq R^* \quad \text{and} \quad f_1 \geq R^* f_0$$

In fact, a slightly stronger condition is given by Gastwirth [6] and Reid [7], i.e.

$$\frac{f_1}{f_0} = R^* + \frac{R^* - 1}{R_U - 1} \frac{1}{f_0} \quad (1)$$

or equivalently that

$$f_1 = R^* f_0 + \frac{R^* - 1}{R_U - 1} \quad (2)$$

In the original application [2], the relative risk of cancer was 5 to 10, so the inequality shows that in order for another factor U to fully explain the smoking–lung cancer association, not only would U need to increase an exposed individual's lung cancer risk five-fold, the proportion of smokers exposed to U needs to be at least five times that of non-smokers. The authors [2] believed that it was not plausible that such an unobserved variable U existed. In the spermicide data, a variable U with $R_U = 2$ and prevalence $f_0 = 0.2$ in the control group would need to have a prevalence of 0.488 in the user group to explain the increased risk. If $R_U = 3$, the required prevalence would be 0.368.

If one has some knowledge of R_U , the effect of the omitted variable, (2) implies that the prevalence of U in the exposed group must exceed $R^* f_0$ by a meaningful amount. If one has a plausible range of values for f_0 , then the required value of f_1 can be determined. Alternatively, if one has prior knowledge of f_0 as the maximum value of $f_1 = 1$, using (2)

one finds that $1 \geq R^* f_0 + (R^* - 1)/(R_U - 1)$, which implies

$$R_U \geq \frac{R^*(1 - f_0)}{1 - R^* f_0} \geq R^* \quad (3)$$

Hence, the minimum relative risk R_U required for another agent to explain the observed R^* exceeds R^* .

While Cornfield *et al.* [2] preferred the relative risk measure for assessing causality, the risk difference is also useful in public health. If U is to entirely explain the observed difference $D^* = P(Y=1|X=1) - P(Y=1|X=0)$, then Y and X are independent for given U . The risk difference associated with an unobserved variable is calculated as $D_U = P(Y=1|U=1) - P(Y=1|U=0)$. Then one must have $(f_1 - f_0)D_U \geq D^*$, and this implies both $D_U \geq D^*$ and $f_1 - f_0 \geq D^*$ [8].

Sensitivity analysis for assessing causality has been discussed extensively by Rosenbaum and Rubin [9] and Rosenbaum and Krieger [10] when the sampling variation is also incorporated. The relationship between sensitivity analysis and causality is thoroughly explored by Rosenbaum [11] and Pearl [12].

3. THE REVERSE CORNFIELD INEQUALITY

The reverse Cornfield inequality enables us to assess whether an unobserved covariate might be masking a meaningful effect. It is especially useful when a study yields a ‘suggestive’ finding, e.g. the observed value of the estimated risk exceeds 1.50 and the p -value is less than 0.20.

In this section, we assume that the model for the response probability is

$$\pi_{X,U} = H(\alpha + \beta X + \gamma U)$$

where $H(\cdot)$ is a twice differentiable function. As in the original inequality, the model assumes there is no interaction between X and U . If the difference between the absolute risks arises from an additive model, $\pi_{X,U} = \alpha + \beta X + \gamma U$. The relative risks are obtained from a multiplicative model, $\log(\pi_{X,U}) = \alpha + \beta X + \gamma U$. The odds ratios are obtained from a logit model $\text{logit}(\pi_{X,U}) = \alpha + \beta X + \gamma U$. With the exception of the proof of the ‘Reverse Cornfield Inequality’ for the multiplicative model, all derivations are presented in the appendix.

3.1. Additive model

For the additive model,

$$\pi_{X,U} = \alpha + \beta X + \gamma U$$

so the relevant risk differences are:

$$D^* = \pi_1 - \pi_0, \quad D_X \equiv \beta = \pi_{1,u} - \pi_{0,u}, \quad D_U \equiv \gamma = \pi_{x,1} - \pi_{x,0}, \quad D_{X,U} = \pi_{1,1} - \pi_{0,0}$$

Here D^* is the observed risk difference, D_X is the true risk difference of the agent X , D_U is the risk difference due to U . Under this model, the omitted variable U interacts additively with the agent X , i.e. $D_{X,U} = D_X + D_U$.

Proposition 3.1

The observed risk difference D^* and true risk difference D_X of the agent are related by

$$D^* = D_U(f_1 - f_0) + D_X \tag{4}$$

In the usual application, f_1 is assumed to exceed f_0 , so D^* overestimates the effect of X . If $f_0 > f_1$, D^* will underestimate the true risk D_X .

3.2. Multiplicative model

Cornfield [2] originally considered a multiplicative model

$$\log(\pi_{X,U}) = \alpha + \beta X + \gamma U \tag{5}$$

and that the true value of $\beta = 0$. Because U was omitted, however, the estimate of β^* from the fitted model $\log(\pi_X) = \alpha^* + \beta^* X$ was greater than 0.

In the multiplicative model (5), the measures of relative risk are:

$$R^* = \frac{\pi_1}{\pi_0}, \quad R_X \equiv e^\beta = \frac{\pi_{1,u}}{\pi_{0,u}}, \quad R_U \equiv e^\gamma = \frac{\pi_{x,1}}{\pi_{x,0}}, \quad R_{X,U} = \frac{\pi_{1,1}}{\pi_{0,0}}$$

The agents X, U act multiplicatively, i.e. $R_{X,U} = R_X R_U$.

Lemma 3.1

Let $a, b, c, d > 0$,

$$g(x) = \frac{ax + b}{cx + d}$$

is an increasing (decreasing) function of x if and only if $ad - bc > (<) 0$.

Proof

$$g'(x) = \frac{ad - bc}{(cx + d)^2} > 0 \quad \text{if and only if } ad - bc > 0$$

Proposition 3.2 (Reverse Cornfield inequality)

If $0 < f_1 < f_0 < 1$ and $R_U \geq 1$, we have

$$1 \leq \frac{R_X}{R^*} \leq \frac{f_0}{f_1} \quad \text{and} \quad \frac{R_X}{R^*} \leq R_U \tag{6}$$

Proof

$$\begin{aligned} R^* &= \frac{\pi_1}{\pi_0} = \frac{\pi_{1,1}P(U=1|X=1) + \pi_{1,0}P(U=0|X=1)}{\pi_{0,1}P(U=1|X=0) + \pi_{0,0}P(U=0|X=0)} \\ &= \frac{\frac{\pi_{1,1}}{\pi_{0,0}}f_1 + \frac{\pi_{1,0}}{\pi_{0,0}}(1-f_1)}{\frac{\pi_{0,1}}{\pi_{0,0}}f_0 + \frac{\pi_{0,0}}{\pi_{0,0}}(1-f_0)} \end{aligned}$$

$$\begin{aligned}
 &= \frac{f_1 R_U R_X + (1 - f_1) R_X}{f_0 R_U + (1 - f_0)} \\
 &= R_X \frac{f_1 R_U + (1 - f_1)}{f_0 R_U + (1 - f_0)} \tag{7}
 \end{aligned}$$

When $f_0 > f_1$, the R.H.S. of (7) is a decreasing function of R_U by Lemma 3.1 and attains its minimum (maximum) when $R_U \rightarrow \infty$ ($R_U = 1$), hence,

$$\frac{f_1}{f_0} < \frac{R^*}{R_X} \leq 1$$

Setting $\tau = R_X/R^*$, the above inequality implies that $\tau f_1 < f_0$. From (7), we obtain

$$R_U = 1 + \frac{\tau - 1}{f_0 - \tau f_1} \geq 1 + \frac{\tau - 1}{f_0} \geq \tau$$

3.3. A logit model

Now assume that the response probability follows a logit model, i.e.

$$\log \left(\frac{\pi_{X,U}}{1 - \pi_{X,U}} \right) = \alpha + \beta X + \gamma U$$

The true odds ratios due to the exposure X and the omitted variable U are

$$\phi_X \equiv OR_{X|U} = \frac{\pi_{1,u}(1 - \pi_{0,u})}{\pi_{0,u}(1 - \pi_{1,u})} = e^\beta, \quad u = 0, 1 \tag{8}$$

$$\phi_U \equiv OR_{U|X} = \frac{\pi_{x,1}(1 - \pi_{x,0})}{\pi_{x,0}(1 - \pi_{x,1})} = e^\gamma, \quad x = 0, 1 \tag{9}$$

where β and γ are the log-odds ratios.

If the variable U is omitted from the analysis, then the observed odds ratio is

$$\phi^* = \frac{\pi_1/(1 - \pi_1)}{\pi_0/(1 - \pi_0)} = e^{\beta^*} \tag{10}$$

where

$$\pi_0 = P(R = 1|X = 0) = \frac{e^{\alpha+\gamma} f_0}{1 + e^{\alpha+\gamma}} + \frac{e^\alpha (1 - f_0)}{1 + e^\alpha} \tag{11}$$

$$\pi_1 = P(R = 1|X = 1) = \frac{e^{\alpha+\beta+\gamma} f_1}{1 + e^{\alpha+\beta+\gamma}} + \frac{e^{\alpha+\beta} (1 - f_1)}{1 + e^{\alpha+\beta}} \tag{12}$$

Proposition 3.3

Let $\underline{\pi}$ be the (baseline) response probability when both X and U equal 0, i.e. $\underline{\pi} = e^\alpha / (1 + e^\alpha)$, and the corresponding odds is $\underline{\phi} = e^\alpha$. The observed odds ratios ϕ^* and the true odds ratio ϕ_X are related by a factor τ ,

$$\tau \equiv \frac{\phi_X}{\phi^*} = \frac{(1 - f_0) + \phi_U(f_0 + \underline{\phi})}{(1 - f_1) + \phi_U(f_1 + \underline{\phi} \cdot \phi_X)} \times \frac{1 + \underline{\phi} \cdot \phi_X[(1 - f_1)\phi_U + f_1]}{1 + \underline{\phi}[(1 - f_0)\phi_U + f_0]} \quad (13)$$

Corollary 3.4

The ratio $\tau = \phi_X / \phi^*$ satisfies the following bounds:

$$\begin{aligned} 1/\phi_U \leq \tau \leq \phi_U & \quad \text{if } \phi_U \geq 1 \\ 1/\phi_U \geq \tau \geq \phi_U & \quad \text{if } \phi_U \leq 1 \end{aligned}$$

Corollary 3.5

Under the rare disease assumption, i.e. $\underline{\pi} \rightarrow 0$ or equivalently $\alpha \rightarrow -\infty$, then the observed odds ratio ϕ^* can be approximated as

$$\phi^* = \phi_X \frac{(1 - f_1) + \phi_U f_1}{(1 - f_0) + \phi_U f_0} \quad (14)$$

Under the rare disease assumption, the odds ratio approximates the relative risk, hence (14) is equivalent to equation (7).

Corollary 3.6

When $\phi_X \geq 1$ and $f \equiv f_0 = f_1$,

$$\phi_X > \phi^*$$

unless $\phi_X = 1$ or $\phi_U = 1$ or $f = 0$ or $f = 1$.

Comment: This result shows that if $\phi_X > 1$, even when U is balanced between the exposed and control groups, the marginal odds ratio ϕ^* is always less than ϕ_X . Similarly, when $\phi_X < 1$, we have $\phi_X < \phi^* < 1$. Gail *et al.* [13] showed that in randomized studies, when the prevalences of relevant variables should be nearly equal in both groups, only the parameters in linear and log-linear models have asymptotically unbiased estimators. Especially, they noted that even in the 'balanced' case, the parameter estimators of a logistic model are biased towards zero. Our results are consistent with their finding.

In the following, we assume that $\phi_X > 1$. Tables II and III present τ , the ratio of ϕ_X to ϕ^* for different values of $\beta (= \log \phi_X)$ and $\gamma (= \log \phi_U)$ when $\underline{\pi} = 0.1$ and 0.5. For example, when $\underline{\pi} = 0.1$, $\phi_X = 2.7$ and $\gamma = -1$, Table II shows that $\tau = 1.014$. When the common prevalence f is near 0 or 1, the ratio τ is close to 1. In both tables, as the effect of exposure ϕ_X increases, the ratio τ increases, i.e. the underestimation bias increases with ϕ_X . The ratio τ also increases as the absolute value of $\log \phi_U$ increases.

Figure 1 presents a three-dimensional plot of the relationship between the relative bias τ and $\log \phi_X, \log \phi_U$ when the prevalence $f = 0.5$ and the baseline response probability $\underline{\pi} = 0.5$.

Table II. The ratio of ϕ_X to ϕ^* when the baseline probability $\pi = 10$ per cent and U is equally prevalent.

π	f	β	$\phi_X = e^\beta$	$\gamma = \log(\phi_U)$							
				-2	-1.5	-1	-0.5	0.5	1	1.5	2
0.1	0.10	0.5	1.6	1.005	1.004	1.002	1.001	1.002	1.009	1.025	1.051
0.1	0.25	0.5	1.6	1.012	1.009	1.006	1.002	1.003	1.017	1.043	1.085
0.1	0.50	0.5	1.6	1.021	1.016	1.009	1.003	1.004	1.019	1.047	1.090
0.1	0.75	0.5	1.6	1.025	1.016	1.008	1.002	1.003	1.012	1.031	1.060
0.1	0.90	0.5	1.6	1.018	1.010	1.005	1.001	1.001	1.006	1.014	1.028
0.1	0.10	1.0	2.7	1.013	1.010	1.006	1.002	1.004	1.020	1.053	1.104
0.1	0.25	1.0	2.7	1.031	1.024	1.014	1.005	1.008	1.037	1.096	1.184
0.1	0.50	1.0	2.7	1.055	1.040	1.022	1.007	1.010	1.044	1.108	1.207
0.1	0.75	1.0	2.7	1.063	1.041	1.021	1.006	1.007	1.029	1.072	1.141
0.1	0.90	1.0	2.7	1.044	1.026	1.011	1.003	1.003	1.013	1.033	1.066
0.1	0.10	1.5	4.5	1.026	1.020	1.012	1.004	1.007	1.032	1.081	1.153
0.1	0.25	1.5	4.5	1.061	1.046	1.027	1.009	1.013	1.061	1.153	1.287
0.1	0.50	1.5	4.5	1.107	1.076	1.041	1.012	1.017	1.074	1.181	1.342
0.1	0.75	1.5	4.5	1.119	1.076	1.037	1.010	1.012	1.051	1.125	1.243
0.1	0.90	1.5	4.5	1.082	1.046	1.021	1.005	1.005	1.023	1.058	1.115

Table III. The ratio of ϕ_X to ϕ^* when the baseline probability $\pi = 50$ per cent and U is equally prevalent.

π	f	β	$\phi_X = e^\beta$	$\gamma = \log(\phi_U)$							
				-2	-1.5	-1	-0.5	0.5	1	1.5	2
0.5	0.10	0.5	1.6	1.031	1.021	1.011	1.003	1.003	1.009	1.016	1.022
0.5	0.25	0.5	1.6	1.067	1.044	1.022	1.006	1.005	1.018	1.035	1.049
0.5	0.50	0.5	1.6	1.098	1.062	1.030	1.008	1.007	1.026	1.052	1.078
0.5	0.75	0.5	1.6	1.089	1.052	1.023	1.006	1.006	1.022	1.047	1.077
0.5	0.90	0.5	1.6	1.051	1.027	1.012	1.003	1.003	1.011	1.026	1.048
0.5	0.10	1.0	2.7	1.074	1.047	1.022	1.006	1.005	1.015	1.027	1.037
0.5	0.25	1.0	2.7	1.155	1.097	1.046	1.011	1.010	1.033	1.060	1.085
0.5	0.50	1.0	2.7	1.219	1.132	1.061	1.015	1.013	1.048	1.093	1.139
0.5	0.75	1.0	2.7	1.188	1.105	1.046	1.011	1.011	1.040	1.086	1.141
0.5	0.90	1.0	2.7	1.102	1.054	1.022	1.005	1.005	1.021	1.049	1.089
0.5	0.10	1.5	4.5	1.126	1.076	1.034	1.008	1.006	1.020	1.035	1.047
0.5	0.25	1.5	4.5	1.260	1.154	1.069	1.016	1.013	1.043	1.078	1.110
0.5	0.50	1.5	4.5	1.354	1.202	1.089	1.021	1.018	1.064	1.124	1.183
0.5	0.75	1.5	4.5	1.286	1.154	1.065	1.016	1.014	1.055	1.116	1.190
0.5	0.90	1.5	4.5	1.148	1.076	1.031	1.007	1.007	1.029	1.067	1.122

Although the ratio τ between the true odds ratio ϕ_X and the marginal odds ratio ϕ^* is small in most situations, it can be substantial. For example, τ can reach 1.992 when $\log(\phi_X) = 2.5$ and $\log(\phi_U) = -2.5$.

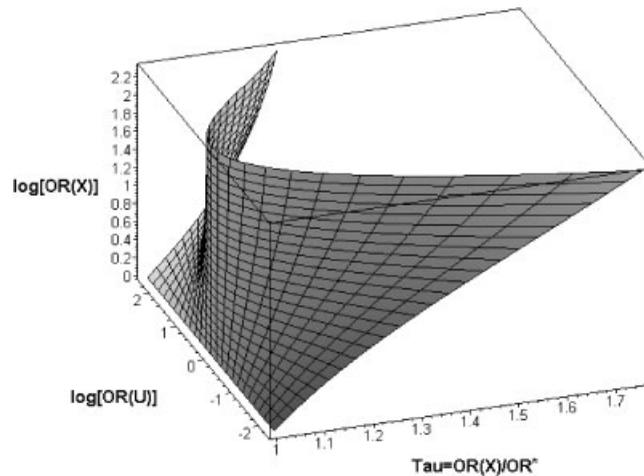


Figure 1. Ratio $\tau = \phi_X/\phi^*$ when $f_0 = f_1 = 0.5$ and $\pi = 0.5$.

4. APPLICATION TO THE STUDY OF THE SPERMICIDE EFFECT ON BIRTH DEFECTS

Motivated by criticism of a legal decision 'Wells vs Ortho Pharmaceutical Corp.', Gastwirth [14] examined the studies discussed in the decision to assess whether the critics, who claimed that the decision and its affirmation by the Court of Appeals showed that the legal system was unable to deal with scientific evidence, was justified. The decision found that limb defects of the baby were results of the mother's exposure to a spermicide after her last menstrual period (LMP) at the time the limb buds were formed. The court also found that there was sufficient evidence from two studies prior to the mother's use of the drug that the defendant should have warned of an increased risk of limb defects. Data from studies cited in the opinion are presented in Reference [6] and the criticisms, subsequent studies and legal opinions are reviewed in Reference [14].

The first published study [15] that suggested spermicide was associated with limb defects ($R_X = 2.98$) also indicated that tranquilizers were a risk factor with a similar relative risk. A later study [16], which apparently was not submitted as evidence in the trial, found a relative risk of 3.7 for tranquilizer use and smoking for all birth defects.

As most of the critics relied on the study of Mills *et al.* [5] that did not find a significant increased risk, we examine its sensitivity to the fact that it did not include data on tranquilizer use by the mothers. Table I presented the basic risk data from this study indicating a relative risk of 1.24. Mills *et al.* noted that in the study population, spermicide users were significantly older, of higher parity and more educated than users of other methods (all $P < 0.0001$). They drank less alcohol ($P < 0.0001$) and smoked fewer cigarettes ($P < 0.0001$).

Although Mills *et al.* [5] did not present the smoking prevalences in both groups, another study in the same era [17] found that 31 per cent of spermicide users past their last menstrual period (LMP) smoked while 39 per cent of non-users did. We assume that the proportion

Table IV. The true relative risk R_X for different values of (R_U, f_0, f_1) when $R^* = 1.241$.

f_0	f_1	R_U	R_X
0.40	0.30	2.0	1.316
0.40	0.30	3.0	1.363
0.40	0.30	4.0	1.395
0.40	0.30	5.0	1.418
0.32	0.30	3.7	1.277
0.34	0.30	3.7	1.314
0.36	0.30	3.7	1.351
0.38	0.30	3.7	1.388
0.40	0.30	3.7	1.425
0.40	0.38	3.7	1.273
0.40	0.36	3.7	1.308
0.40	0.34	3.7	1.345
0.40	0.32	3.7	1.384
0.40	0.30	3.7	1.425

of tranquilizer use in the smoking group equals that of the non-smoking group and will use these results as a basis of our sensitivity analysis.

Because the joint use of tranquilizer and cigarettes U is omitted from the original analysis, we only observed the marginal relative risk $R^* = \pi_1/\pi_0$ and estimated it by $\hat{R}^* = b/n_1/a/n_0$ in Table I. By the Reverse Cornfield Inequality (7), the (observed) marginal relative risk R^* and the relative risk R_X of spermicide are related by a factor τ , i.e. $R_X = \tau R^*$, where

$$\tau = \frac{f_0 R_U + (1 - f_0)}{f_1 R_U + (1 - f_1)} \quad (15)$$

Table IV lists the true values of the relative risk R_X for different values of (R_U, f_0, f_1) . We see that when $f_0 > f_1$, the true relative risk R_X exceed the observed relative risk 1.241 and the difference between R_X and R^* increases with R_U . When $f_0 = 0.40$, $f_1 = 0.30$ and $R_U = 3.7$, the true relative risk R_X would be 1.425.

For testing whether spermicides are teratogenic (cause a birth defect), i.e.

$$H_0: R_X = 1 \quad \text{vs} \quad H_1: R_X > 1 \quad (16)$$

the test statistic is based on observed relative risk \hat{R}^* , i.e. $T = \log(\hat{R}^*)$. If we assume that $\bar{\pi} = \pi_0 = \pi_1$ under the null, the rejection region is

$$T = \log(\hat{R}^*) > Z_{1-\alpha} \sigma_0$$

where $(\sigma_0)^2 = (1 - \bar{\pi})/\bar{\pi}((1/n_0) + (1/n_1))$ and $\bar{\pi}$ can be estimated by $\bar{p} = (a + b)/(n_0 + n_1)$. From Table I, the observed relative risk \hat{R}^* is 1.241, the p -value for a one-sided test is 0.11. This closely agrees with the two-sided p -value of 0.21 reported in Table III by Mills *et al.* [5]. For a one-sided (two-sided) test with α level 0.05, the rejection region is $\hat{R}^* > 1.305$ ($|\hat{R}^*| > 1.373$).

Because of the imbalance in smoking between the spermicide users and non-users, under $H_0: R_X = 1$, π_1 no longer equals π_0 . Indeed, (15) implies that when $f_0 > f_1$ and $R_U > 1$, the expected value of the observed relative risk $R^* = 1/\tau < 1$. For convenience, $1/\tau$ will be denoted by k . The imbalance of U also affects the variance of the null distribution of \hat{R}^* , which is given by the following proposition.

Proposition 4.1

Under the null, the true variance of $\log(\hat{R}^*)$ is

$$(\sigma_0^*)^2 = V^*(\log(\hat{R}^*)|H_0) = \frac{1 - \pi_0}{n_0 \pi_0} + \frac{1 - k\pi_0}{n_1 k \pi_0} \quad (17)$$

where π_0 can be estimated by maximizing the product-binomial likelihood.

Since the variable U was omitted from the analysis, the test based on $\log(\hat{R}^*)$ is not unbiased. From (15), $\log(R^*)$ and $\log(R_X)$ are related by

$$\eta = \log(R^*) - \log(R_X) = \log(k) = \log \left\{ \frac{f_1 R_U + (1 - f_1)}{f_0 R_U + (1 - f_0)} \right\}$$

Under $H_0: R_X = 1$, the test statistic $T = \log(\hat{R}^*) \sim N(\eta, (\sigma_0^*)^2)$. For the observed relative risk $r = 1.24$, the true p -value adjusting for the imbalance of U between the groups is

$$p = P(T > \log(r)|H_0) = P \left(\frac{\log(\hat{R}^*) - \eta}{\sigma_0^*} > \frac{\log(r) - \eta}{\sigma_0^*} \right) = 1 - \Phi \left(\frac{\log(r) - \eta}{\sigma_0^*} \right)$$

The change in the p -values for different values of the relative risk of U and different prevalences of U in the exposed and control groups are displayed in Figure 2. The initial values for the three parameters are $R_U = 3.7$, $f_0 = 39$ per cent and $f_1 = 31$ per cent. In the sensitivity analysis, we change the value of one parameter at a time. From Figure 2(a), we see that as the effect of U , R_U , increases, the p -value decreases. If $R_U = 5$, the one-sided p -value declines to 0.02. In Figure 2(b), when $f_1 = 0.31$, as f_0 increases from 0.31 to 0.41, the p -value decreases from its original value 0.11 to 0.018 when $R_U = 3.7$. Figure 2(c) shows a similar pattern as the prevalence of U in the exposed group declines.

The above sensitivity analysis makes the unrealistic assumption that all smokers used tranquilizers. As other studies reported that spermicide users tended to drink less as well as smoke less than non-users, they probably were more health conscious than non-users. Thus, it is reasonable to assume that a smaller fraction of spermicide users who smoked used tranquilizers too. In Table V we assume that between one-fourth to one-half of the smokers used tranquilizers in the control group but the probability that a spermicide user took tranquilizers was less than that of the control group. Notice that in most cases an imbalance in the prevalence of smoking and tranquilizer use could have created a sufficient underestimate of the risk of spermicide use to mask a statistically significant result if their joint risk exceeds 3.0.

The size α of the test also will be affected by the imbalance in U . The true significance level of the test is

$$\alpha^* = P(T > Z_{1-\alpha} \sigma_0 | H_0) = P \left(\frac{\log(\hat{R}^*) - \eta}{\sigma_0^*} > \frac{Z_{1-\alpha} \sigma_0 - \eta}{\sigma_0^*} \right) = 1 - \Phi \left(\frac{Z_{1-\alpha} \sigma_0 - \eta}{\sigma_0^*} \right)$$

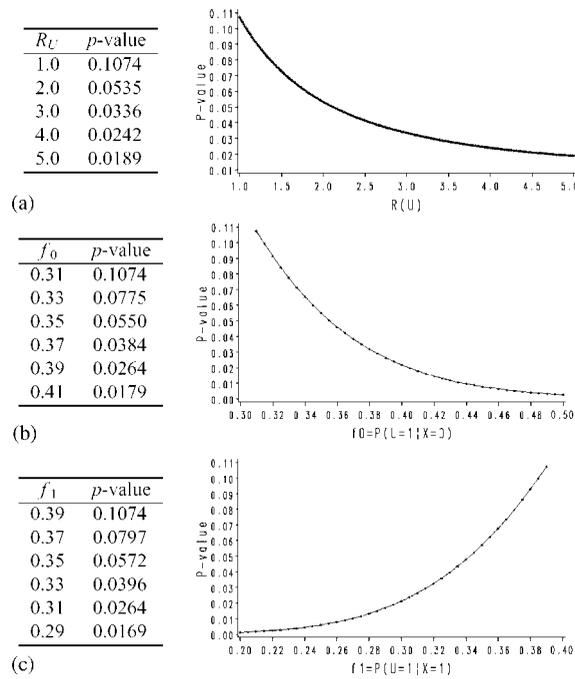


Figure 2. The true p -value of the test $H_0:R_X = 1$ vs $H_1:R_X > 1$ (Initial values $R_U = 3.7, f_0 = 39$ per cent, $f_1 = 31$ per cent): (a) R_U changes; (b) f_0 changes; and (c) f_1 changes.

Table V. The true relative risk R_X for different values of (R_U, f_0, f_1) when a smaller percentage of exposed individuals have the omitted risk factor.

R_X value		R_U			
f_0	f_1	2	3	4	5
0.20	0.15	1.294	1.335	1.368	1.395
0.20	0.10	1.353	1.447	1.526	1.594
0.20	0.05	1.417	1.578	1.725	1.860
0.15	0.10	1.296	1.343	1.383	1.417
0.15	0.075	1.327	1.402	1.468	1.526
0.10	0.075	1.269	1.294	1.316	1.335
0.10	0.05	1.299	1.353	1.402	1.447

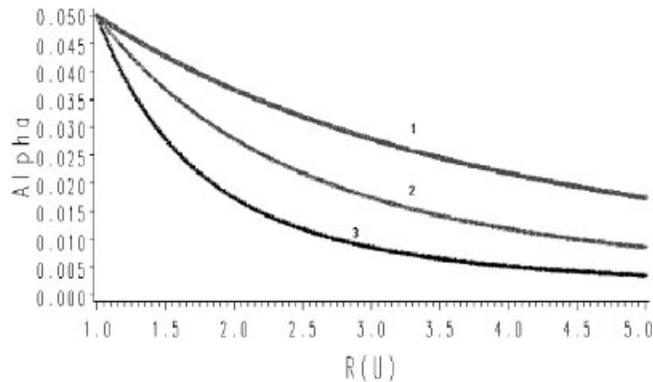


Figure 3. Significance level of the test $H_0:R_X=1$ vs $H_1:R_X>1$ (1: $f_0=0.1, f_1=0.075$; 2: $f_0=0.2, f_1=0.1$; 3: $f_0=0.4, f_1=0.3$).

Figure 3 plots the significance level as a function of R_U for different prevalence pairs. The lowest curve (3) shows the α level for $f_0=0.4, f_1=0.3$. The upper two curves show the α levels when only one-half or one-quarter of the smokers use tranquilizers. We can see that the significance level of the test is *less* than 0.05 when the omitted variable U is more prevalent in the control group. Thus, the original study actually used a lower value of α than 0.05. As the effect of U increases, the significance level α of the test decreases. The α level decreases faster when $f_0=0.4$ and $f_1=0.3$ because the difference between f_0 and f_1 is larger than in the other cases, $f_0=0.2, f_1=0.15$ or $f_0=0.1, f_1=0.075$. When $R_U=3.7$, all the one-sided α -levels are below 0.03.

Also, under the alternative $H_1:R_X>1$, the true power of the test will change to

$$1 - \beta^* = P(T \leq Z_{1-\alpha}\sigma_0|H_1) = 1 - \Phi\left(\frac{Z_{1-\alpha} - (\log(R_X) - \eta)/\sigma_0}{\sigma_1/\sigma_0}\right) \tag{18}$$

The variance of $\log(\hat{R}^*)$ under the alternative is given by

$$(\sigma_1)^2 = V(\log(\hat{R}^*)|H_1) = \frac{1 - \pi_0}{n_0\pi_0} + \frac{1 - \pi_1}{n_1\pi_1}$$

where the estimates are $\hat{p}i_0$ and $\hat{p}i_1 = b/n_1$.

Figure 4 presents the power of the test under the same conditions as in Figure 3. The original power of the test under the alternative $H_1:R_X=2$ is 0.996. It decreases to 0.945 when $R_U=4.5$ and $f_0=0.4, f_1=0.3$. The power under the alternative $H_1:R_X=1.5$ decreases much more, from 0.81 to 0.50. Hence, omitting the confounder U makes the test invalid as the α is lower than 0.05. Moreover, the power decreases too. In Table VI, we present the true significance level and power of the test for $R_U=3.7$ under different prevalences f_0 and f_1 .

It is interesting to explore the implication of our sensitivity analysis of the non-significant result reported by Mills *et al.* [5]. When the omitted variable U has an R_U of 3.7, the α level is compromised. When $f_0=0.4$ and $f_1=0.3$, a one-sided 0.05 level test has actual level

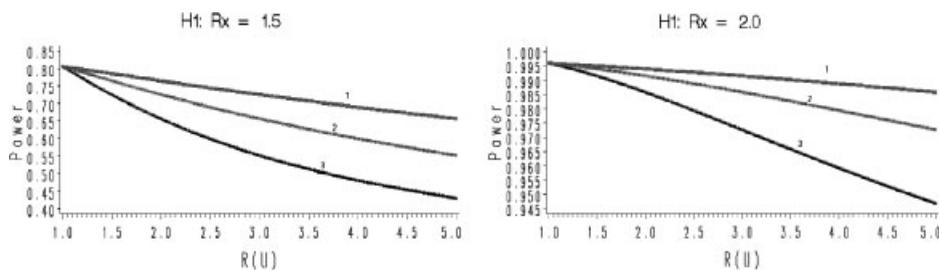


Figure 4. The true power of the test $H_0:R_X = 1$ vs $H_1:R_X > 1$ (1: $f_0 = 0.1, f_1 = 0.075$; 2: $f_0 = 0.2, f_1 = 0.15$; 3: $f_0 = 0.4, f_1 = 0.3$).

Table VI. The effect of an omitted variable with $R_U = 3.7$ on the α level of the test $H_0:R_X = 1$ vs $H_1:R_X > 1$ and the power under alternatives $H_1:R_X = 1.5$ and $H_1:R_X = 2.0$.

f_0	f_1	α level	Power	
			$H_1:R_X = 1.5$	$H_1:R_X = 2.0$
0.4	0.40	0.050	0.807	0.996
0.4	0.35	0.019	0.673	0.987
0.4	0.30	0.006	0.501	0.963
0.4	0.25	0.001	0.316	0.905
0.4	0.20	0.000	0.158	0.784
0.2	0.20	0.050	0.807	0.996
0.2	0.15	0.013	0.616	0.981
0.2	0.10	0.002	0.370	0.927
0.2	0.05	0.000	0.151	0.776
0.1	0.10	0.050	0.807	0.996
0.1	0.08	0.028	0.724	0.991
0.1	0.06	0.014	0.623	0.982
0.1	0.04	0.006	0.507	0.965
0.1	0.02	0.003	0.385	0.933

0.006 and if $f_0 = 0.2, f_1 = 0.15$, the actual level is 0.013. Similar results hold for two-sided tests.

While the effect on the power of a study to detect a relative risk of 2.0 is not severely compromised, it does fall below the original value (greater than) 0.99 given by Mills *et al.* [5]. Courts have accepted relative risk as low as 1.55 [18]. Under the alternative $H_1:R_X = 1.5$, the power of the test falls from the original value 0.807 when the omitted factor is balanced in the two groups to less than 0.600 when $f_1 < f_0/2$ for the prevalences in Table VI.

If $R_U = 3.7$ and $f_0 = 0.39, f_1 = 0.31$, Figure 2(b) yields a two-sided p -value of 0.053. If R_U were only 2.6, the two-sided p -value would rise to 0.1. If $f_0 = 0.2, f_1 = 0.15$ and $R_U = 3.7$, the

two-sided p -value would be 0.12. All these p -values are not far from significance. Moreover, there is controversy in the literature as to the effect of smoking itself on limb defects and the study by Smith *et al.* [15] also indicated that a threatened abortion during pregnancy is a risk factor. Unfortunately no information about the relative prevalences of this risk in the user and non-user groups were reported in the Mills *et al.* [5] study, so we could not examine its potential impact. Thus, our analysis reinforces the recommendation of Cordero and Layde [19] that the relationship between spermicide use and birth defects deserves further studies. These studies should incorporate tranquilizer use, smoking pattern and other risk factors in order to reach a sound conclusion.

5. DISCUSSION

The use of the reverse Cornfield inequality allows statisticians to examine the potential effect of omitted variables on non-significant findings when those risk factors are likely to be more prevalent in the control group. The analysis of the spermicide study illustrates that in this situation, an omitted factor that increases the risk under study leads to a test with less than the prespecified significance level α with a consequent decrease in power. This makes it more difficult to uncover a true risk factor. In practice, this suggests that the data underlying non-significant but suggestive results ($p < 0.20$) should be carefully examined to answer whether other risk factors could be more prevalent in the control group.

APPENDIX

Proof of Proposition 3.1

Note that

$$\pi_1 = P(Y=1|X=1) = \pi_{1,1}f_1 + \pi_{1,0}(1-f_1)$$

$$\pi_0 = P(Y=1|X=0) = \pi_{0,1}f_0 + \pi_{0,0}(1-f_0)$$

and

$$D_X = \pi_{1,0} - \pi_{0,0} = \pi_{1,1} - \pi_{0,1}$$

$$D_U = \pi_{0,1} - \pi_{0,0} = \pi_{1,1} - \pi_{1,0}$$

Hence the observed risk difference

$$\begin{aligned} D^* &= \{(\pi_{1,1} - \pi_{0,0})f_1 + (\pi_{1,0} - \pi_{0,0})(1-f_1)\} - \{(\pi_{0,1} - \pi_{0,0})f_0 + (\pi_{0,0} - \pi_{0,0})(1-f_0)\} \\ &= D_{X,U}f_1 + D_X(1-f_1) - D_Uf_0 \\ &= (D_X + D_U)f_1 + D_X(1-f_1) - D_Uf_0 \\ &= D_X + D_U(f_1 - f_0) \end{aligned}$$

Proof of Proposition 3.3

By (11) and (12),

$$\frac{\pi_0}{1 - \pi_0} = \frac{\frac{e^{\alpha+\gamma}}{1 + e^{\alpha+\gamma}}f_0 + \frac{e^\alpha}{1 + e^\alpha}(1 - f_0)}{\frac{1}{1 + e^{\alpha+\gamma}}f_0 + \frac{1}{1 + e^\alpha}(1 - f_0)}$$

$$\frac{\pi_1}{1 - \pi_1} = \frac{\frac{e^{\alpha+\beta+\gamma}}{1 + e^{\alpha+\beta+\gamma}}f_1 + \frac{e^{\alpha+\beta}}{1 + e^{\alpha+\beta}}(1 - f_1)}{\frac{1}{1 + e^{\alpha+\beta+\gamma}}f_1 + \frac{1}{1 + e^{\alpha+\beta}}(1 - f_1)}$$

so that

$$\frac{\pi_0}{1 - \pi_0} = e^\alpha \times \frac{e^\gamma(1 + e^\alpha)f_0 + (1 + e^{\alpha+\gamma})(1 - f_0)}{(1 + e^\alpha)f_0 + (1 + e^{\alpha+\gamma})(1 - f_0)}$$

$$= e^\alpha \times \frac{(1 - f_0) + e^\gamma(f_0 + e^\alpha)}{1 + e^\alpha[(1 - f_0)e^\gamma + f_0]}$$

$$\frac{\pi_1}{1 - \pi_1} = e^{\alpha+\beta} \times \frac{e^\gamma(1 + e^{\alpha+\beta})f_1 + (1 + e^{\alpha+\beta+\gamma})(1 - f_1)}{(1 + e^{\alpha+\beta})f_1 + (1 + e^{\alpha+\beta+\gamma})(1 - f_1)}$$

$$= e^{\alpha+\beta} \times \frac{(1 - f_1) + e^\gamma(f_1 + e^{\alpha+\beta})}{1 + e^{\alpha+\beta}[(1 - f_1)e^\gamma + f_1]}$$

Substituting $\pi_1/(1 - \pi_1)$ and $\pi_0/(1 - \pi_0)$ into (10), we obtain

$$\phi^* = \phi_X \times \frac{(1 - f_1) + e^\gamma(f_1 + e^{\alpha+\beta})}{(1 - f_0) + e^\gamma(f_0 + e^\alpha)} \times \frac{1 + e^\alpha[(1 - f_0)e^\gamma + f_0]}{1 + e^{\alpha+\beta}[(1 - f_1)e^\gamma + f_1]} \quad (\text{A1})$$

Rewriting (A1) in terms of the odds ratios ϕ_X, ϕ_U and $\underline{\phi}$ yields

$$\tau \equiv \frac{\phi_X}{\phi^*} = \frac{(1 - f_0) + \phi_U(f_0 + \underline{\phi})}{(1 - f_1) + \phi_U(f_1 + \underline{\phi} \cdot \phi_X)} \times \frac{1 + \underline{\phi} \cdot \phi_X[(1 - f_1)\phi_U + f_1]}{1 + \underline{\phi}[(1 - f_0)\phi_U + f_0]}$$

Proof of Corollary 3.5

Let

$$g(\xi, f) = \frac{(1 - f) + e^\gamma(f + e^\xi)}{1 + e^\xi[(1 - f)e^\gamma + f]} = \frac{1 + e^{\xi+\gamma} + (e^\gamma - 1)f}{1 + e^\xi[1 - (e^\gamma - 1)f]}$$

Then by (A1),

$$\phi^* = \phi_X \frac{g(\alpha + \beta, f_1)}{g(\alpha, f_0)}$$

If $\phi_U = e^\gamma \geq 1$, $g(\xi, f)$ is an increasing function of f . Thus, ϕ^* attains its maximum when $f_1 = 1$ and $f_0 = 0$, and the value is $\phi_X \cdot \phi_U$ by (A1). It attains the minimum ϕ_X/ϕ_U when $f_1 = 0$ and $f_0 = 1$.

When $\phi_U = e^\gamma \leq 1$, $g(\xi, f)$ is a decreasing function of f . ϕ^* attains its minimum value, $\phi_X \cdot \phi_U$, at $f_1 = 0, f_0 = 1$, and its maximum ϕ_X/ϕ_U at $f_1 = 1, f_0 = 0$.

Proof of Corollary 3.5

When $\alpha \rightarrow -\infty$, the baseline odd $\underline{\phi} = e^\alpha \rightarrow 0$. From equation (13),

$$\tau = \frac{\phi_X}{\phi^*} = \frac{(1 - f_0) + \phi_U f_0}{(1 - f_1) + \phi_U f_1}$$

Proof of Corollary 3.6

When $f \equiv f_0 = f_1$, then $\tau = \phi_X/\phi^* = A/B$, where

$$A = [(1 - f) + \phi_U(f + e^\alpha)] \times [1 + e^\alpha \phi_X((1 - f)\phi_U + f)]$$

$$B = [(1 - f) + \phi_U(f + e^\alpha \phi_X)] \times [1 + e^\alpha((1 - f)\phi_U + f)]$$

Because

$$A - B = \phi_U e^\alpha + (1 - f + \phi_U f)e^\alpha \phi_X [(1 - f)\phi_U + f]$$

$$- \phi_U e^\alpha \phi_X + (1 - f + \phi_U f)e^\alpha [(1 - f)\phi_U + f]$$

$$= e^\alpha (\phi_X - 1) \{ (1 - f + \phi_U f) [(1 - f)\phi_U + f] - \phi_U \}$$

$$= e^\alpha (\phi_X - 1) (\phi_U - 1)^2 f (1 - f) \geq 0 \tag{A2}$$

Hence $\tau = \phi_X/\phi^* = A/B \geq 1$.

Proof of Proposition 4.1

The unconditional likelihood of the two proportions (π_0, π_1) in a 2×2 table is

$$L_u(\pi_0, \pi_1) = B(a; n_0, \pi_0) B(b; n_1, \pi_1) = \binom{n_0}{a} \pi_0^a (1 - \pi_0)^{n_0 - a} \binom{n_1}{b} \pi_1^b (1 - \pi_1)^{n_1 - b} \tag{A3}$$

Let $\hat{R}^* = (b/n_1)/(a/n_0)$. To improve the applicability of the normal approximation $\log(\hat{R}^*)$ is usually used.

Because $a \sim \text{Bin}(n_0, \pi_0)$ and $b \sim \text{Bin}(n_1, \pi_1)$ are independent, by Central Limit Theorem, $p_0 = a/n_0 \sim N(\pi_0, \pi_0(1 - \pi_0)/n_0)$ and $p_1 = b/n_1 \sim N(\pi_1, \pi_1(1 - \pi_1)/n_1)$. As $\log(\hat{R}^*) = \log(p_1) - \log(p_0)$, $V(\log(\hat{R}^*)) = V(\log(p_1)) + V(\log(p_0))$. By the δ -method, the variance of $\log(p_0)$ is given by

$$V(\log(p_0)) \approx \left[\frac{d \log(\pi)}{d\pi} \Big|_{\pi=\pi_0} \right]^2 V(p_0) = \left[\frac{1}{\pi_0^2} \right] \frac{\pi_0(1 - \pi_0)}{n_0} = \frac{1 - \pi_0}{n_0 \pi_0}$$

Likewise, the variance $V(\log(p_1)) \approx (1 - \pi_1)/(n_1 \pi_1)$, hence, the variance of $\log(\hat{R}^*)$ is $(1 - \pi_0)/n_0 \pi_0 + (1 - \pi_1)/n_1 \pi_1$.

If we assume $\pi_0 = \pi_1 = \pi_c$, π_c is estimated by $\bar{p} = (a + b)/(n_0 + n_1)$, then $\hat{V}(\log(\hat{R}^*)) = (1/n_0 + 1/n_1)\bar{p}/(1 - \bar{p})$. Under the alternative, π_0 and π_1 are independent and are estimated by $p_0 = a/n_0$ and $p_1 = b/n_1$. Hence $\hat{V}(\log(\hat{R}^*)) = (1 - p_0)/n_0 p_0 + (1 - p_1)/n_1 p_1$. If a confounding risk factor U is omitted, from (7), $k = \pi_1/\pi_0 \neq 0$ under the null $H_0: R_X = 1$. The variance of $\log(\hat{R}^*)$ is $(1 - \pi_0)/n_0 \pi_0 + (1 - k\pi_0)/kn_1 \pi_0$, where the MLE $\hat{\pi}_0$ of π_0 is the solution to equation

$$\frac{d \log L_u(\pi_0, k\pi_0)}{d\pi_0} = \frac{a}{\pi_0} - \frac{n_0 - a}{1 - \pi_0} + \frac{kb}{k\pi_0} - \frac{k(n_1 - b)}{1 - k\pi_0} = 0$$

This reduces quadratic equation

$$g(\pi_0) = k(n_0 + n_1)\pi_0^2 - [k(a + n_1) + (b + n_0)]\pi_0 + (a + b) = 0$$

In our analysis $k = \pi_1/\pi_0 < 1$, hence $g(0) = a + b > 0$ and $g(1) = (k - 1)(n_0 - a) < 0$. There is a unique solution $\hat{\pi}_0 \in (0, 1)$.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. Justine Shults for her helpful discussions and insightful comments on the subject during the preparation of this manuscript. We also wish to thank Dr George Reed for explaining the power calculation in the 1982 study and two referees and the editor for many helpful comments. This research was supported in part by an NSF grant.

REFERENCES

1. Greenhouse SW. Jerome Cornfield's contributions to epidemiology. *Biometrics* 1982; **38S**:33–46.
2. Cornfield J, Haenszel W, Lilienfeld A, Shimkin M, Wynder E. Smoking and lung cancer. *Journal of National Cancer Institute* 1959; **22**:173–203.
3. Gastwirth JL, Greenhouse SW. Biostatistical concepts and methods in the legal setting. *Statistics in Medicine* 1995; **14**:1641–1653.
4. Li YP, Proport KJ, Rosenbaum PR. Balanced risk set matching. *Journal of the American Statistical Association* 2001; **96**:870–882.
5. Mills J, Harley EE, Reed GF, Berendes HW. Are spermicides teratogenic? *Journal of the American Medical Association* 1982; **248**:2148–2151.
6. Gastwirth JL. *Statistical Reasoning in Law and Public Policy*. Academic: New York, 1988.
7. Reid N. Cornfield's Lemma. In *Encyclopedia of Statistical Sciences*, Kotz S, Read CB, Banks DL (eds). Wiley: New York, 1996; 163–164.
8. Gastwirth JL, Krieger AM, Rosenbaum PL. Cornfield's inequality. In *Encyclopedia of Biostatistics*, Armitage P, Colton T (eds). Wiley: New York, 1998; 952–955.
9. Rosenbaum PL, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society, Series B* 1983; **45**:212–218.
10. Rosenbaum PL, Krieger A. Sensitivity analysis for two-sample permutation inferences in observational studies. *Journal of the American Statistical Association* 1990; **85**:493–498.
11. Rosenbaum PL. *Observational Studies* (2nd edn). Springer: New York, 2002.
12. Pearl J. *Causality: Models, Reasoning and Inference*. Cambridge University Press: New York, 2000.
13. Gail M, Wieand S, Piantadosi S. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and missing covariates. *Biometrika* 1984; **71**:431–444.
14. Gastwirth JL. The role of epidemiological evidence in product liability cases: a reexamination of *Wells v. Ortho and Key Pharmaceuticals*. Manuscript, 2001.

15. Smith ESO, Dafoe CS, Miller JR, Banister P. An epidemiological study of congenital reduction deformities of the limbs. *British Journal of Preventive and Social Medicine* 1977; **31**:39–41.
16. Bracken MB, Holford TR. Exposure to prescribed drugs in pregnancy and association with congenital malformations. *Obstetrics and Gynecology* 1981; **58**:336–344.
17. Polednak AP, Janerich DT, Glebatis DM. Birth weight and birth defects in relation to maternal spermicide use. *Teratology* 1982; **26**:27–38.
18. Gastwirth JL. Epidemiology as legal evidence. In *Encyclopedia of Epidemiologic Methods*, Gail MH, Benichou J (eds), vol. 160. Wiley: New York, 2000; 386–390, 289–303.
19. Cordero JF, Layde PM. Vaginal spermicides, chromosomal abnormalities and limb reduction defects. *Family Planning Perspective* 1983; **15**:16–18.