Perspectives on Weak Associations in the Health Sciences: An Epistemological Note

Daniele Chiffi & Dario Gregori

(University of Padova)

Abstract

Weak statistical associations in the health sciences can be a potential clue to causation under certain conditions. The present paper aims at pointing out some methodologies capable to investigate the possible causal import of weak statistical associations. Debiasing and prior elicitation techniques are presented in order to possibly overcome some problems placed on the edge between statistics and causality. We draw attention to the choice of the statistical techniques in handling weak associations, as every model brings with it certain (tacit) philosophical assumptions of methodological nature.

ملخّص

يمكن للترابطات الإحصائية الضعيفة فى علوم الصحة أن تقدم مفتاحاً ممكناً لحل لغز السببية تحت شروط معينة . تهدف المقالة الحالية إلى الإشارة لبعض المنهجيات القادرة على فحص الفحوى السببية الممكنة للترابطات الإحصائية . وسوف تُقدم تقنيات فض التحامل والتوضيح المتقدم من أجل تجاوز بعض المشكلات الموجودة فى المنطقة الواقعة بين الإحصاء والعِلية ، مع لفت الانتباه إلى عملية اختيار التقنيات الإحصائية فى معالجة الترابطات الضعيفة ، لأن كل نموذج يستحضر معه بعض الافتراضات الفلسفية (المضمرة) التى لها طابع منهجى.

Résumé

Les faibles associations statistiques dans les sciences médicales peuvent fournir une clé potentielle pour résoudre, sous certaines conditions, le probléme de la causalité. L'article présent tend à examiner quelques unes des méthodologies susceptibles d'investir le possible import causal des faibles associations statistiques. Des tehniques de dépolarisation et de première élucidation sont présentées dans le but de traiter dans les termes du possible quelques problèmes placés entre les statistiques et la causalité. Vu que chaque modèle apporte avec lui quelques assomptions philosophiques (tacites) de nature méthodologique, nous attirerons l'attention au choix des techniques statistiques dans le traitement des faibles associations.

Introduction

Probabilistic theory of causation covers a non-secondary but recent role in the health sciences because of the increasing contribute of statistics in medicine. Within the framework of probabilistic causation, in particular in etiological epidemiology and clinical epidemiology for the effect of novel treatments, like in oncology (Pless and Weinberg 2011), there is the issue of "weak associations", namely those associations between

factors which are at most twice greater in the exposes if compared to the control group and which might be a possible clue to causality. It is quite common to name those associations between risk factors and outcomes characterized by a relative risk (namely, which provides the ratio of the probability – incidence in epidemiologic terms – of the disease in the exposed population to that in the unexposed population) , less than 2 (e.g., the association between smoking and myocardial infarction) as "weak associations" (Boffetta 2010). Indeed, if the ratio between the incidence in the exposure group and in the control group is greater than 2, this means that an individual in the exposure group is *more likely than not* to develop the disease. If the relative risk is 1.5, the component 1 is the background risk while only the component 0.5 is due to a particular exposure, while if relative risk is 2.5, the background risk is 1 and the risk-component due to exposure is 1.5. In such a case, an individual in the exposure group is *more likely than not* to develop the disease, since the risk-component associated to the exposure is greater than the background risk-component.

The acknowledgment of a 'true' weak association is a key epidemiological notion and "an inferential challenge" (Weed 1997). On the one hand, weak associations can be easily affected by confounding and bias, but, on the other hand, the possibility that a weak association can be a causal relationship cannot be ruled out a priori. More generally, from a public health perspective, the fact that relative risk is < 2 does not entail that such risk can be neglected if the exposure prevalence of the risk factor leads to a high Levinson's population attributable risk, which is the portion of outcomes which could be avoided if the risk factor exposure would be removed.

This paper aims at outlining the reasons for adopting an eclectic view on probabilistic causality in order to integrate epidemiologic evidence with the individual level of clinical knowledge in the context of medical research and practice.

The next section is devoted to the issue of weak associations, while Section 2 deals with the assessment of causal inference in clinical epidemiology and its relevance for applied philosophy. Section 3 is devoted to analyse certain limits of the probabilistic approach to causality and to outline the methodology of prior elicitation which is intended to soundly convert experts' information into probabilistic information. Some concluding remarks in Section 4 underline the necessity of integrating epidemiologic with clinical evidence.

1. Weak associations

In epidemiologic research, Rothman and Poole have suggested the possibility of a 'strengthening programme' for weak associations obtained by reducing misclassification and bias and by focusing on population at low risk in order to avoid the interaction between study factors and other possible causes (Rothman and Poole 1988). For instance, imagine that we want to study the relation between a weak risk factor, say (*D*) drinking alcohol, and lung cancer (*LC*). As it is a well establish fact that there is a very strong association between LC and smoking cigarettes (*SCs*), consequently if we want to study the relationship between *W* and *LC*, we must exclude smokers from the target population, otherwise the majority of cases would be due to *SCs*. In this way the weak probabilistic association might eventually indicate a possible causal relationship (Szklo and Nieto 2004; Cornfield, Haenszel et al. 2009), even if other risk factors cannot be excluded. Similar ideas have been also presented in philosophy of science, see (Salmon 2006).

Let us consider now, as a way of example, the associations that are detached in nutritional epidemiology, where weak statistical associations are very common. In this field it is quite common to find studies which report a relative risk estimate of 0.8-1.2, namely a weak association. Higher estimates of relative risk may be due to bias and errors. Nonetheless, the evidence of weak association in nutritional studies, on the one hand, may entail some important concerns from the public health perspective since the population attributable risk is high because of the high prevalence of exposure in the population. On the other hand, an association which is acknowledged weak at population level can conclude to be an important element of causality at individual level. Moreover, it has been recently observed that weak associations in nutritional epidemiology can be enhanced and better understood epidemiologic studies are conducted on different populations trying to replicate the results (Van Staveren, Burema et al. 1988) as well integrating different forms of evidence. When data are presented in a study it may be difficult to assess the real association between the factors of interests and that is why it is necessary to present scientific findings in a way that may avoid the rise of paradoxes. Thus, the framing of data and epidemiologic information is an essential part of the work of the researcher together with the explication of the causal structure that he (or she) is trying to prove. Nevertheless, dealing with statistical weak associations among populations from an epidemiologic perspective does not suffice to detect a causal pathway in an individual case. The knowledge of the epidemiological frequencies in a population might be silent with respect to a single causal event since it is necessary to

know other individual conditions – based on patient's conditions – that may shape the causal process. Such individual conditions can be probabilistically estimated by means of an epistemic account of probability based on the rational beliefs in a causal hypothesis about a single event in which *the conditions determined by other possible known causal factors that may have contributed to the occurrence of the effect are at least partially eliminated*. This eclectic view on the combinations of probabilities is, for instance, supported in (Hacking 2001). The validity and limits of such a view will be evaluated in Section 3. The next section offers a methodological outlook of causal inference in epidemiology.

2. Causal inference in clinical epidemiology

The discovery of a new causal relationship between a risk factor and a certain outcome is always a big step for clinical knowledge, even if only in a few cases there is a strong causal relation between a condition – more likely a single genetic condition in case of Mendelian (monogenetic) diseases – and a resulting outcome. In any case, such association is not deterministic, since one has to take into account the level of *penetrance*, that is the probability of a phenotype given the genotype, and *expressivity*, which is the variation in a phenotype among persons given a particular genotype. Moreover, the influence of environmental and epigenetic conditions cannot be ruled out for many diseases. Hence, even in genetic based diseases there is a set of different (and in certain cases unknown) risk factors which contribute to a statistical association with a specific outcome.

It has been remarked that two main standpoints appears to be involved in the determination of a causal relationship in the biomedical sciences: the *probabilistic* and the (biological, chemical, genetic, social, etc.) *mechanistic* dimensions of causality, whose features are often difficult to be matched, but it does not seem plausible to conceive them separately (Broadbent 2011; Campaner 2011). In any case, other standpoints on causality are non-marginal in the health sciences: e.g., in randomized controlled trials causal claims are based on *counterfactual* conditions between the treatment group and the control group. A very same individual cannot be tested in both groups, therefore it is assumed because of randomization that if an individual of the control group would have undergone the treatment, then he would have been exposed to the same risks and benefits of the other individuals in the treatment group and this is a counterfactual assumption. Furthermore, it has been observed that once an association is detached by means of a randomized controlled trial, there is no warrant that such association could be

externally valid outside the study population (Cartwright 2007; Thompson 2010), even if this is a common feature of all types of epidemiologic studies.

It is worth noting that the debate on causality in clinical epidemiology mainly affects the determination of the nature of the epidemiological research itself. The dispute on the nature of clinical epidemiology, in fact, ranges from positions where clinical epidemiology is conceived as a "black-box" intended to discover the strength of association between efficacy of treatments od the presence of adverse risk factors and outcome (Savitz 1994; Greenland, Gago-Dominguez et al. 2004) up to others assuming that clinical epidemiology must "open the box" in order to understand the causal relationships by means of the findings and theories of other biological, social, chemical and genetic sciences (Skrabanek 1994; Weed 1998). For sure, the black-box strategy involves a higher level of abstraction which can fail in understanding the web of causation since such strategy can merely detect associations; nevertheless, such high level of abstraction in the black-box strategy may suggest new risk factors for which there is not yet, for instance, a biological explanation. In any case, the probabilistic account of causation is required by both tendencies in clinical epidemiology and that is why it has such a central role. In any case, there is also no universal consent, from an epidemiological point of view, on what makes a distinction between a causal relationship and a mere statistical interaction. To this purpose Bradford Hill proposed that causality shows some aspects which may considered as indicators of a possible causal relationship. In the past, in fact, there was the tradition of interpreting Bradford Hill's considerations on some aspects of statistical association as causal criteria (Hill 1965; Ward 2009). These aspects of causation concern the *strength of association*, the *consistency* of association that is the repeated observation of the association on different populations in different contexts, the *specificity* of association for which a cause can only lead to a single effect, the *temporal relationship* for which a cause has to precede the occurrence of the effect, *biological plausibility*, *coherence* for which "our data should not seriously conflict with the general known facts of the natural history and biology", *experimental evidence* and *analogy*. As observed by Rothman and Greenland, if Hill's considerations on causality are conceived as causal criteria, then many counterexamples can invalidate their universal validity in the clinical sciences (Rothman and Greenland 2005). However, Richard Doll in a *Fisher memorial lecture* clarified that only temporal relationship is *conditio sine qua non*, since exposure is an essential requirement for causality in the health domain (Doll 2002). Moreover, it can be worth noting that these causal associations are handled by Hill's

aspects at the group level, leaving without indication the causal relationship at individual level. Causality, instead, is essentially connected with the choice (or the interplay) of an individual or a population level of risk. In any case, causality at individual level can be hardly conclusively established, but it possible to assess – from a probabilistic point of view – if an individual will develop more likely than not a disease because of an exposure to a risk factor if some of the following conditions are fulfilled (Lagiou, Adami et al. 2005):

1) The exposure is an established cause of the disease.
2) The exposure of the individual has to be similar (duration, intensity, latency) to the exposure causing the disease.
3) The disease of the individual must be similar to that which is aetiologically associated to the exposure.
4) The individual has not to be exposed to other risk factors.
5) The relative risk must be greater than 2.

Notice that this perspective on causality makes particularly sense when one wants to acknowledge (at individual level) already known causes, while it seems less adequate for the discovery of new causes. At any rate, condition 5 is quite problematic for our discussion. In fact, such condition appears to partially rule out the possibility of establishing a causal relationship in presence of weak associations and this seems to be unlikely. In case of weak associations, we suggest to integrate population-based probabilistic data with the epistemic probability based on the medical knowledge of the clinician. Such an integrative view of probabilistic causality has some advantages since it possible to deal with weak association in probabilistic causality but there is the well-known problem of assessing clinician's probability attribution. Anyway, some new procedures of statistical expert elicitation can facilitate this task. We will go further in this discussion in the next section, after first presented some paradoxes connected with the notion of probabilistic causality.

3. Probabilistic Causality and some paradoxes

The probabilistic account of causality is *prima facie* based on the intuitive idea that a condition φ causes an event ψ if $p(\psi|\varphi) > (\psi|\neg\varphi)$ together with other conditions. It has been observed, though, that causality seems to be an asymmetrical relation, while this is not the case for probability relation and effects seem never (or almost never) to occur

before their causes. On the other hand, a probabilistic account of causality can be connected with the concept of risk (which is a key notion in medicine), for risk is the *probability* of an uncertain outcome in connection with the magnitude of the effect and the consequences. Unfortunately, many biases are associated to probabilistic causality. Some of them, which are extremely important in the health sciences, are the following:

1) *Simpson's paradox* occurs when the probabilistic association between two variables is inverted in each subpopulation of a target population since an undetected risk factor is causally related to the outcome and associated with the determinant (Yule 1903; Simpson 1951; Heydtmann 2002). Such paradox, which is involved in the comparison between two groups which are different according to a determinant of health like age, sex etc.., can be overcome by a standardisation or the addition of new relevant (causal) information in order to avoid confounding. A variable which fulfils the following conditions can be considered as a confounder if:

i) it is associated with the risk factor in the population.
ii) it is be related to the disease, excluding the relation between risk factors and disease.
iii) it is not a link in the causal pathway to disease (Hannan 1996).

As a way of example, assume that in relation to disease *X* the hospital *A* has 56% of positive outcomes and the hospital *B* 58%. Indeed it seems that *B* presents better results rather than A. Now, imagine that 88% of the *A* cases are related to serious diseases (*SDs*) and there is a positive outcome (*PO*) in 50% of these cases (namely in 44% of the *A* cases), while the residual 12% are mild cases (*MCs*) with 100% of positive outcomes (namely, 12% of the *A* cases). The hospital *B*, instead, has 36% of (*MCs*) of which 90% shows a *PO* (32.4% of the *B* cases), while the remaining 64% of *SD* shows a *PO* in 40% (namely 25.6% of the *B* cases). Therefore, it is better to undergo the treatment in the hospital *A* in any case, since if I have a serious disease then the hospital *A* has 50% of a *PO* while the hospital *B* has 40% and if I have a mild disease then the probability of a *PO* is 100% in hospital *A* and 90% in the hospital *B*. By contrast, abstracting from the severity of the disease, the hospital *B* must be preferred (Table 1). Hence, the clinician has to understand if the addition of new *individual* clinical information can reverse the effect of the exposure in order to not violate condition (4).

2) *Berkson's paradox* is connected with a selection bias for which persons have different probabilities of being included in the study sample with respect to the pertinent features to be analysed in the study (Berkson 1946; Roberts, Spitzer et al. 1978; Armstrong 1998; Sadetzki, Bensal et al. 2003). If a patient has two diseases, then his (or her) probability of being hospitalised is greater than the probability connected to either disease individually. For instance, in case-control studies the recruitment of the population among hospitalized cases and controls may face this paradox, since people with multiple diseases can be over-represented in the contingency tables regarding hospitalization data, while this is not the case at the overall population level.

3) *Neyman Bias*. When there is a gap between the exposure and the selection it is possible to incur a *prevalence-incidence* bias (also known as *Neyman bias* (Neyman 1955)), consisting in including prevalent cases in the case-control studies. In such studies an association may be spurious if the risk factor affects, for instance, survival. A supposed association detected by hospital records between myocardial infarction and snow shovelling may be affected by bias due to the possibility that some people may have died during the driveways, thus not succeeding in arriving at the hospital. As a consequence, the level of association between a risk factor and an outcome can be easily underrated, because of the inadequacy of the case group. Thus, disease duration can involve the modification of the assessment of the association between determinants and an outcome when such association is based on prevalent cases. Sackett clarifies that "a late look at those exposed (or affected) early will miss fatal and other short episodes, plus mild or silent cases and cases in which evidence of exposure disappears with disease onset" (Sackett 1979). Note that also for this bias the knowledge of the causal structure of a process plays a central role.

4) *Lindley-Jeffreys paradox*. In any circumstances the acknowledgment of a probabilistic association may depend on the choice of the perspective on probability adopted by the researcher and on the knowledge of different sources of available evidence. For instance, it can be the case that one adopts a statistical hypothesis testing which may indicate that the null hypothesis H_0 must be rejected according to a specific P value, while from a Bayesian perspective H_0 is more probable than the alternative hypothesis H_1 and this seems absurd. This condition is named *Lindley-Jeffreys paradox* and states that a statistical significant small association H_1 may be less likely than H_0 by calculating the Bayes factor (B) which expresses the ratio of the odds of prior and posterior information that can

favour one hypothesis rather than the other (Jeffreys 1939; Lindley 1957; Ioannidis 2008).

Being a *ratio*, *B* can assume values between 0 and infinite. If $B>1$, then H_0 receives more probabilistic support rather than H_1, if $B<1$ then H_1 receives more probabilistic support rather than H_0 and if $B=1$, then H_0 and H_1 have the same probabilistic support. Therefore, a significant association which has been acknowledged by means of classical statistical hypothesis testing can be disconfirmed through Bayesian methods. If $B=1/100$, this fact suffices to convert a prior probability of 0.9 in the truth of the null hypothesis into a probability of 0.08 and in this case there is very strong evidence against the null hypothesis, while e.g. if the initial probability in the null hypothesis is 0.9, a Bayes factor of 1/10 gives a poster probability of $H_0=0.47$. Thus H_0 is less likely than not to be true. More generally, a Bayes factor of 1/5 expresses a *weak evidence*, a Bayes factor of 1/10 a *moderate evidence*, a Bayes factor of 1/20 *a substantial strength of evidence* and a Bayes factor of 1/100 shows a *very strong strength of evidence* between probabilistic factors (Goodman 1999).

The importance of the choice of the probabilistic model has become evident in connection with the association between vitamin E supplementation and an increase of all-cause mortality which was detached by classical meta-analyses (Miller, Pastor-Barriuso et al. 2005). By contrast, a Bayesian meta-analysis of the same articles has showed no association between vitamin E supplementation and an increase of all-cause mortality (Berry, Wathen et al. 2009). Hence, the choice of a model entails the implicit acceptance of some methodological and causal assumptions which may lead toward different findings and interpretations. In any case, a meta-analysis can improve the acknowledgment of a causal relationship even if it is not its primary role. Note that *Lindley-Jeffreys paradox* can receive an adequate interpretation if it is considered being related to the possibility of integrating statistical population data with the assessment of clinical (individual) knowledge.

The aforementioned biases show the unreliability of the probabilistic causality *per se*, notably in the clinical setting. That is why expert's clinical opinions can improve medical judgment integrating population based probabilistic data with clinical knowledge by means of procedures of expert's elicitation together with a Bayesian framework. The elicitation of probabilities of expert's beliefs can be an important tool when assessing causality in the clinical context. There are many techniques to elicit probabilities (Slottje, van der Sluijs et al. 2008). According to one of these techniques, the expert is asked to evaluate whether the actual value of a quantity is higher or lower than a certain number.

This can be carried out, for instance, by means of graphical tools such as probabilistic wheels. Alternatively, the expert is asked to fix the value of a quantity such that the probability of higher or lower values turns out to be some specific amount. We do not go into the analysis of these techniques. What is important for our discussion is that the clinician should be aware of the possibility of incurring inconsistencies and also of the availability of some methods for avoiding these.

More generally, it is a well-known fact that intuitive reasoning involving probabilistic computations is affected by many heuristics and cognitive biases which influence both experts and lay people (Tversky and Kahneman 1974; Kahneman and Tversky 1979). For instance, the framing of the information plays an essential role in decision making, e.g. people are more likely to undergo a treatment if it is communicated to them that there is 80% of no counter indication in place of communicating that there is 20% of a negative outcome related to the treatment. Thus, the knowledge of cognitive heuristics and biases in connection with "*debiasing techniques*" can be a good tool for expert elicitation in order to assess the probabilistic measure of expert's knowledge and, thus, integrating clinical information with scientific evidence.

Final Remarks

We have discussed the complexity of dealing with the statistical and causal structure of weak associations. We have observed that it is important to frame the data of a study in a way that can minimize the occurrence of some paradoxes and cognitive bias and this also requires a specification of a causal mechanism when it is possible, especially when the association is weak. The choice of a whatsoever model entails the tacit acceptance of some assumptions. For instance, the choice of assuming or not a "black box strategy" is a methodological constraint that might partially modify the epidemiologic models and, sometimes, the interpretation of the findings of a study. Another methodological assumption that we have presented is the distinction between a general and an individual level of causation, specifically when dealing with weak associations. Sound strategies of experts' prior elicitation can integrate population-based data with the judgment of the clinician in order to mitigate the impact of the paradoxes associated to the notion of probabilistic causality, which occur at population level in the health sciences, with clinician's knowledge associated with current clinical practice.

References

Armstrong, B. G. (1998). "Effect of measurement error on epidemiological studies of environmental and occupational exposures." Occup Environ Med **55**(10): 651-656.

Berkson, J. (1946). "Limitations of the Application of Fourfold Table Analysis to Hospital Data." Biometrics Bulletin **2**(3): 47-53.

Berry, D., J. K. Wathen, et al. (2009). "Bayesian model averaging in meta-analysis: vitamin E supplementation and mortality." Clinical Trials **6**(1): 28-41.

Boffetta, P. (2010). "Causation in the Presence of Weak Associations." Critical Reviews in Food Science and Nutrition **50**: 13-16.

Broadbent, A. (2011). Inferring causation in epidemiology: mechanisma, black boxes, and contrasts. Causality in the Sciences. P. M. Illari, F. Russo and J. Williamson. Oxford, Oxford University Press**:** 45-69.

Campaner, R. (2011). "Understanding mechanisms in the health sciences." Theoretical Medicine and Bioethics **32**(1): 5-17.

Cartwright, N. (2007). "Are RCTs the Gold Standard?" BioSocieties **2**(01): 11-20.

Cornfield, J., W. Haenszel, et al. (2009). "Smoking and lung cancer: recent evidence and a discussion of some questions." International Journal of Epidemiology **38**(5): 1175-1191.

Doll, R. (2002). "Proof of causality - deduction from epidemiological observation." Perspectives in Biology and Medicine **45**(4): 499-515.

Goodman, S. N. (1999). "Toward Evidence-Based Medical Statistics. 2: The Bayes Factor." Annals of Internal Medicine **130**(12): 1005-1013.

Greenland, S., M. Gago-Dominguez, et al. (2004). "The value of risk-factor ("black-box") epidemiology." Epidemiology (Cambridge, Mass.) **15**(5): 529-535.

Hacking, I. (2001). An introduction to probability and inductive logic. Cambridge, U.K. ; New York, Cambridge University Press.

Hannan, M. T. (1996). "Is it a risk factor or confounder? A discussion of selected analytic methods using education as an example." Arthritis & Rheumatism **9**(5): 413-418.

Heydtmann, M. (2002). "The nature of truth: Simpson's Paradox and the limits of statistical data." QJM **95**(4): 247-249.

Hill, A. B. (1965). "The Environment and Disease: Association or Causation?" Proc R Soc Med **58**: 295-300.

Ioannidis, J. P. A. (2008). "Effect of Formal Statistical Significance on the Credibility of Observational Associations." American Journal of Epidemiology **168**(4): 374-383.

Jeffreys, H. (1939). Theory of probability. Oxford,, The Clarendon press.

Kahneman, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." Econometrica **47**(2): 263-291.

Lagiou, P., H.-O. Adami, et al. (2005). "Causality in cancer epidemiology." European Journal of Epidemiology **20**(7): 565-574.

Lindley, D. V. (1957). "A Statistical Paradox." Biometrika **44**(1-2): 187-192.

Miller, E. R., R. Pastor-Barriuso, et al. (2005). "Meta-Analysis: High-Dosage Vitamin E Supplementation May Increase All-Cause Mortality." Annals of Internal Medicine **142**(1): 37-46.

Neyman, J. (1955). "Statistics--Servant of All Sciences." Science **122**(3166): 401-406.

Pless, M. and U. Weinberg (2011). "Tumor treating fields: concept, evidence and future." Expert Opin Investig Drugs **20**(8): 1099-1106.

Roberts, R. S., W. O. Spitzer, et al. (1978). "An empirical demonstration of Berkson's bias." J Chronic Dis **31**(2): 119-128.

Rothman, K. J. and S. Greenland (2005). "Causation and causal inference in epidemiology." Am J Public Health **95 Suppl 1**: S144-150.

Rothman, K. J. and C. Poole (1988). "A strengthening programme for weak associations." Int J Epidemiol **17**(4): 955-959.

Sackett, D. L. (1979). "Bias in analytic research." J Chronic Dis **32**(1-2): 51-63.

Sadetzki, S., D. Bensal, et al. (2003). "The limitations of using hospital controls in cancer etiology - one more example for Berkson's bias." European Journal of Epidemiology **18**(12): 1127-1131.

Salmon, W. C. (2006). Four decades of scientific explanation, University of Pittsburgh press.

Savitz, D. A. (1994). "In Defense of Black Box Epidemiology." Epidemiology **5**(5): 550-552.

Simpson, E. H. (1951). "The Interpretation of Interaction in Contingency Tables." Journal of the Royal Statistical Society. Series B (Methodological) **13**(2): 238-241.

Skrabanek, P. (1994). "The Emptiness of the Black Box." Epidemiology **5**(5): 553-555.

Slottje, P., J. van der Sluijs, et al. (2008). Expert Elicitation: Methodological suggestions for its use in environmental health impact assessments. RIV letter report 630004001/2008.

Szklo, M. and F. J. Nieto (2004). Epidemiology: beyond the basics. Sudbury, Mass., Jones and Bartlett.

Thompson, R. P. (2010). "Causality, mathematical models and statistical association: dismantling evidence-based medicine." Journal of Evaluation in Clinical Practice **16**(2): 267-275.

Tversky, A. and D. Kahneman (1974). "Judgment under Uncertainty: Heuristics and Biases." Science **185**(4157): 1124-1131.

Van Staveren, W. A., J. Burema, et al. (1988). "Weak associations in nutritional epidemiology: the importance of replication of observations on individuals." Int J Epidemiol **17**(4): 964-969.

Ward, A. C. (2009). "The role of causal criteria in causal inferences: Bradford Hill's "aspects of association"." Epidemiol Perspect Innov **6**: 2.

Weed, D. L. (1997). "On the use of causal criteria." International Journal of Epidemiology **26**(6): 1137-1141.

Weed, D. L. (1998). "Beyond black box epidemiology." Am J Public Health **88**(1): 12-14.

Yule, G. U. (1903). "Notes on The Theory of Association of Attributes in Statistics " Biometrika **2**(2): 121-134.

	Hospital A		**Hospital B**	
	Mild Cases	Severe Cases	Mild Cases	Severe Cases
Distribution of cases per hospital	12%	88%	36%	64%
Percentage of positive outcomes	100%	50%	90%	40%
Percentage of positive outcomes in relation to the severity of disease	12%	44%	32.4%	25.6%
Aggregated Percentages	56%		58%	

Table 1. An example of Simpson's Paradox in a clinical