Eur J Epidemiol
DOI 10.1007/s10654-014-9928-x

COMMENTARY

On the essentials of etiological research for preventive medicine

Igor Karp · Olli S. Miettinen

Received: 4 June 2014 / Accepted: 10 June 2014

Epidemiological research on the population level (rather than in laboratories) is, ideally at least, intended to advance the knowledge-base of the practice of population-level preventive medicine. To this end it typically addresses etiology of illness, so as to identify risk factors through which prevention could be accomplished.

Despite the accumulated extent of such research, its methodological essentials remain in need of critical consideration. They can be deduced from the essence of etiology as a particular genre of causation in medicine; and very notably, conformity with these essentials is not a feature of 'case–control' nor 'cohort' studies on etiology.

We here draw attention to the essentials of etiological studies for preventive medicine. This we supplement with deduction from these of the essence of confounding specific to this genre of causal research.

I. Karp (✉)
Department of Social and Preventive Medicine, School of Public Health, University of Montreal, Montreal, QC, Canada
e-mail: igor.karp@gmail.com; igor.karp.1@umontreal.ca

I. Karp
Health Risks Division, Centre de recherche du CHUM, University of Montreal, Montreal, QC, Canada

O. S. Miettinen
Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada

O. S. Miettinen
Department of Medicine, Weill Medical College, Cornell University, New York, NY, USA

The essence of etiology

In the dictionary of epidemiology sponsored by the International Epidemiological Association [1], etiology is defined as:

Literally, the science of causes, causality; in common usage, cause.

A related definition in this dictionary is that of etiological study: "A study that aims to unveil causal relationships."

These definitions can be taken to mean that etiology is the object of all causal thought in preventive medicine and in the research underpinning it; they do not make it clear that etiology is the object of only one of the two fundamental types of such causal thought. In books on the theory of causality-oriented epidemiological research, etiology commonly is left undefined, and even the term may not appear.

Thought about a cause-effect relationship addresses what empirically is an antecedent-subsequent relationship. In thinking about etiology in the practice of preventive medicine the subsequent is a given, and the causal question is about the antecedent: Was this subsequent (which is a given) caused by the antecedent in question? This question really focuses on situations in which the antecedent actually was present. With this focus, the question—on the orientational/qualitative level—is, Would the subsequent have occurred but for the antecedent having been there (ceteris paribus)?

In preventive medicine, etiological questions are addressed when the thinking is about causal explanation—inherently *retrospective*—of an extant status of the client's health (in clinical medicine the presence of a particular illness, in community medicine the level of morbidity from a particular illness). In this thinking, the concern is about the causal origin—*etiogenesis*—of the status quo.

Published online: 21 June 2014

The essence of etiological studies

Epidemiological studies on etiology of illness through the first half of the 20th century were characterized by *singularity* of their essential nature. Consistent with the retrospective essence of etiology, each expressly etiological study centered on a group of persons with the illness in question, with the members documented in respect to the potentially causal antecedent in question. This group was coupled with a control group of persons without this illness, similarly documented. These two groups were compared in respect to their respective distributions by the antecedent at issue. In the context of the singularity of the methodological essence of etiological studies at that time, no descriptive adjective was applied to denote the overall design of these studies. Illustrative of that type of study is one published by R. Doll and A. B. Hill in 1950 [2], concerning smoking in the etiology of lung cancer.

With a good deal of this type of research on that etiological topic already completed, Doll and Hill came to the view that "Further … studies of that same kind would seem to us unlikely to advance our knowledge materially or to throw any new light upon the nature of the association," adding that "If, too, there were any undetected flaw in the evidence that such studies have produced, it would be exposed only by some entirely new approach" [3]. On this basis, they proceeded to introduce a novel type of study, one that would "determine the frequency with which the disease appeared, in the future, among groups of persons whose smoking habits were already known" [3].

Having thus introduced a methodological duality into their etiological studies, Doll and Hill for the first time characterized the previous type of study as "retrospective," and the novel type, correspondingly, as "prospective" [3]. This duality, conceptual and linguistic, was eminent in the 1964 Smoking and Health report [4] of the committee advisory on this topic to the Surgeon General of the US: each of the studies on smoking in the etiology of lung cancer was classified as either "retrospective" or "prospective." And in etiological research in general this conceptual duality has been maintained up to the present, though with some evolution in the two concepts, and with the terms changed.

While this duality of types of etiological study did not arise as a perceived implication of the very concept of etiology, the methodological implications of the distinctive essence of etiological causation (above) have been critically considered. This has not led to that duality but to the methodological essentials in this research [5–8] different from those that used to prevail before the 'prospective' study was introduced. This new singularity—it can be viewed as a modern version of the 'retrospective' study—involves four essential elements:

1. Study base, representing the domain designed for the etiological object of the study.
2. Case series, representing all of the instances of the illness event in question occurring in the study base, with these instances (person-moments) documented in respect to the causal contrast at issue (together with whatever else is seen to be relevant).
3. Base series, representing the study base and documented analogously to the case series.
4. Synthesis of the data on these two series to a result on the object of study.

Even in the framework of the prevailing duality in the types of etiological study, the conceptual developments alluded to above, while not uniformly established, can be seen to represent steps toward this singularity. In the context of 'retrospective'/'case–control' studies it has become rather commonplace to perceive the need for a defined study base and to think of both the 'case group' and 'control group' in terms of having the study base as their shared referent. And the rather common deployment of Cox regression in 'prospective'/'cohort' studies is tantamount to identification of the cases occurring in the population-time of the cohort's follow-up and sampling of the person-moments constituting this population-time.

Attainment of the essentials

Given the essentials of any etiological study (above), they should not be seen to result from the design of the study; they should be seen to be requisite a priori features of the study.

By contrast, the process by which these essentials are attained involves many options. An outstanding exception to this stems from the nature of the study base in any etiological study. For obtaining the case and base series, a two-stage process is required: in the first stage, the two series are obtained from the selected, defined source base; and in the second stage these series are reduced to instances from the actual study base (in which the person-moments satisfy the study object's domain criteria, i.a.).

The need for this *two-stage process* in etiological studies is not being appreciated at present; even the very concept of source base is typically absent from contemporary texts on the theory of epidemiological research on etiology. In 'case–control' studies the cases are treated as though deriving directly from the study base itself. And in 'cohort' studies, the population-time of the cohort's follow-up is not treated as constituting merely the source base (for the first-stage case and base series).

Validity of the process

In etiological studies, a triad of possible biases is commonly thought of—'selection,' 'information,' and 'confounding' bias. Of particular note among these is the nature of confoundedness of the study base in these studies: it has to do with the distributions of person-moments of occurrence/nonoccurrence of the health event at issue in the follow-up of a dynamic (open-to-exit) population. But according to the IEA-sponsored dictionary of epidemiology [1], confounding (implicitly, in reference also to etiological studies) has to with distributions at "baseline," even though in a dynamic study population there is no "baseline" for the members' follow-up.

Epilogue

For full appreciation of the requisite features of etiological studies it may be helpful to take note of the counterparts of these in the nonetiological genre of causal research for preventive medicine.

In this nonetiological research, causal thinking about the antecedent-subsequent relationship is conditional on the antecedent, and the causal question is, Will this antecedent (actual or hypothetical), invoked in lieu of its defined alternative, have the effect in question (on the subsequent)? Thus, the causal thinking in this research is about the consequence(s) –*prospective*—of the antecedent at issue.

The essential elements of nonetiological causal studies involve, centrally, the enrolment of a series of entries into prospective presence of the antecedent at issue and of its alternative, and follow-up of the persons involved in respect to the subsequent occurrence/nonoccurrence of the health event at issue. Thus, in these studies, the study population is operationally formed from its source population, and it is a *cohort*. It is in these studies that confounding of the study base can be a matter of distributions at the "baseline" (zero time) of the follow-up.

Such studies for preventive medicine are epitomized by randomized trials on the effects of vaccinations. While commonly considered to be paradigmatic for etiological studies, these trials are, as has been seen here and pointed out before [7], not only unhelpful but actually misleading when used in this role.

References

1. Porta M (Editor), Greenland S, Last JM (Associate Editors). A Dictionary of Epidemiology. 5th ed. Oxford: Oxford University Press; 2008.
2. Doll R, Hill AB. Smoking and carcinoma of the lung: preliminary report. Br Med J. 1950;2:739–48.
3. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits: a preliminary report. Br Med J. 1954;1:1451–5.
4. United States Public Health Service. Smoking and health: report of the advisory committee to the surgeon general of the public health service. Washington: U.S. Department of Health, Education, and Welfare; 1964.
5. Miettinen OS. Estimability and estimation in case-referent studies. Am J Epidemiol. 1976;103:226–35.
6. Miettinen OS. Etiologic research: needed revisions of concepts and principles. Scand J Work Environ Health. 1999;25:484–90.
7. Miettinen OS. Etiologic study vis-à-vis intervention study. Eur J Epidemiol. 2010;25:671–5.
8. Miettinen OS, Karp I. Epidemiological research: an introduction. Dordrecht: Springer; 2012.