MIETTINEN 1989

J Clin Epidemiol Vol. 42, No. 6, pp. 491–496, 1989
Printed in Great Britain.

0895-4356/89 $3.00 + 0.00

THE CLINICAL TRIAL AS A PARADIGM FOR EPIDEMIOLOGIC RESEARCH

OLLI S. MIETTINEN*

Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec, Canada
and Free University, Amsterdam, The Netherlands

Abstract—The extent to which the clinical trial serves, and fails, as a paradigm for epidemiologic research in general is examined. It is argued, first, that the traditional paradigms—investigating epidemic and endemic occurrence of illness in the context of public-health activities, inclusive of the deployment of census, vital and morbidity statistics and sample surveys—are misleading for scientific research. Major examples of the consequences of these paradigms are the preoccupations with time and place, and with "the general population" or some other "target population"—both alien from the vantage of clinical trials and, indeed, of science in general. Then it is shown, by the use of the clinical trial paradigm, that traditional epidemiologic thought and practice in cause–effect research are misguided in the context of such common contexts as the use of empirical contrasts between exposure and unspecified nonexposure, the employment of "representative" distributions of determinants, and, even, as to the belief that cohort and "case–control" studies constitute alternatives to each other. On the other hand, it is argued that for etiologic research the ordinary (parallel) clinical trial is misleading as a paradigm, especially as for learning about the essential temporal aspects of the cause–effect relation.

Clinical trials Epidemiologic methods

INTRODUCTION

In my teaching of theoretical epidemiology [1], the clinical trial has become a very important point of reference. For some central purposes it serves as the supreme paradigm, and for others, I need to caution my students about its inadequacies as a model. As for other teachers and authors, it is my impression that some [2, *inter alia*] have a tendency to make too much of it as a paradigm for nonexperimental epidemiologic research, while others [3, *inter alia*] fail to draw certain important lessons from the theory and practice of clinical trials. The issues are important and subtle, and the variation in views calls for an attempt at a systematic exposition of them.

As is well known, clinical trials are experimental studies on the effects of interventive agents (typically drugs) or interventions (treatments) themselves in clinical medicine. Regardless of whether the conceptual focus is on effects of interventive agents (explanatory trial [4]) or of treatments themselves (pragmatic trial [4]), empirically a clinical trial relates the occurrence of some outcome phenomenon to the categories of a treatment—conditionally on various extraneous determinants of the occurrence, or potential confounders. The interest in occurrence is directed, inherently, to the frequencies of occurrence of various categories of the phenomenon at issue, even if in the context of a quantitative phenomenon the frequency distribution may be characterized by some overall descriptive parameter, such as the mean. The relation of the outcome parameter to the determinant (treatment), so long as it is conditional on the entire set of potential confounders, unknown as well as known, can be given a causal interpretation. The requisite conditionality is

*All correspondence should be addressed to: Dr O. S. Miettinen, McGill University, 1020 Pine Avenue West, Montreal, Quebec, Canada H3A 1A2.
This work was supported by the Institute for Research in Extramural Medicine, Free University, Amsterdam.
This paper is based on an invited lecture to the International Society of Clinical Biostatistics, Gothenburg, 1987.

pursued by randomization, possibly augmented by blocking and/or control in the analysis of the data.

The essence of epidemiologic research has remained less well understood. In recent years, I have come to the view and conviction that such research should be defined as *occurrence research* [1, p. vii], that is, research in which the occurrence of a phenomenon is related to some (potential) determinant(s) of that occurrence—commonly conditionally on some other determinants (potential confounders) [1, pp. 12–16].

These definitions imply that clinical trials are epidemiologic studies. This, coupled with the relatively advanced state of development of the theory and practice of clinical trials in comparison with epidemiologic research at large, implies the importance of seeking to understand the extent to which the clinical trial serves as a paradigm for nonexperimental epidemiologic research.

TRADITIONAL PARADIGMS

For a perspective on the role of the clinical trial as a paradigm for nonexperimental epidemiologic research, cognizance of the traditional paradigms is relevant. A general—and authoritative—sense of these may be gleaned from Lilienfeld's **Foundations of Epidemiology** [3].

With concern for occurrence of illness and related phenomena in "general" human populations, and with particular concern for etiology and prevention on the community level, the traditional epidemiologist relates rates of occurrence to time, place and personal characteristics; and the latter, while subsuming demographic, social, economic, behavioral and biologic ones, does not include clinical actions [3, pp. 1–2]. Given this traditional separation of epidemiology from clinical medicine, it is no wonder that the clinical trial does not, to the traditionalist, have any particular status as a paradigm for epidemiologic research at large.

Outstanding among the traditional paradigms of epidemiologic research is the investigation of epidemics, and a particularly eminent single investigation in this genre is John Snow's investigation of cholera epidemics in London in 1848–1854 [3, pp. 24–25]. (In witness to this, a picture of the epidemiologically famous Broad Street water pump adorns the cover of Lilienfeld's book, *inter alia*.)

Another major tradition that today's epidemiology still draws from is that of census, vital, and morbidity statistics on a national scale [3, pp. 22–23, Chaps 4–5], supplemented by morbidity surveys [3, pp. 115–116].

With deference to these traditions, the traditionalist takes the broad strategy of epidemiologic (etiologic) inquiry to consist of two phases: first, relating occurrence rates in "general" populations to characteristics of those populations in the aggregate, and second, relating individual health outcomes to characteristics of individuals [3, pp. 13–15]. (These two types of inquiry have been termed "descriptive" and "analytic" epidemiology, respectively—in defiance, or ignorance, of general principles of science. Properly, the "analytic method" in science is one in which complex phenomena are "resolved into more simple concepts or general principles so that there will be a transition from the particular to the general" [5]; and in any case, "descriptive" and "analytic" do not properly denote alternatives to each other, because meaningful description requires "penetrating analysis" [6]. On the other hand, descriptive *problems* contrast with ones that are inferential as to causation [1, pp. 11–12].)

The success of John Snow, and others, in applying "the epidemiologic method" to etiologic problems led to the emergence of epidemiology as a "study" [3, p. 1] defined by its concern for occurrence of illness and related phenomena in human populations in the community—with special reference to etiology but without any restriction as to subject-matter (as to the type of health phenomenon or etiologic agent), and with a strong commitment to the methodologic traditions delineated above. Even though concern for etiologic, or occurrence, aspects of phenomena cannot properly define a science, on account of the diversity of subject-matter (cf. morphology), epidemiology—in the subject-matter sense—is viewed by the traditionalist as a field, and even as a science. This outlook is manifest in professorships, societies, journals etc. devoted to "epidemiology". (That the methodology in epidemiologic research be both scientific and singular is not, philosophically, a proper basis for regarding epidemiology as a science [7].)

The methodologic traditions of epidemiology, apart from having given rise to a malformed field of "study" (in the subject-matter sense), are themselves quite inadequate for the purposes of etiologic research itself, to say nothing about

the other realms of occurrence research. Some of these shortcomings will come to focus in examining, in the sections below, the utility of the clinical trial as a paradigm for occurrence research.

TIME AND PLACE

"The epidemiologist is interested in the occurrence of disease by time, place and persons" [3, p. 1], even though such particularism of interest (spatio-temporal specificity) is generally alien in the context of clinical trials and, indeed, in science in general [7]. In clinical trials, traditional epidemiologists' contrary belief [3, pp. 223–224] notwithstanding, there usually is no particularistic "target population" of interest—no population to be "sampled" and to be treated as the referent, or target, of sample-to-population inference. Instead, clinical trials usually address abstract (nonparticularistic) questions—referring to abstract domains, defined not by time and place but by type (as to indication, age and other person-characteristics) —in the spirit of science in general [7].

What the clinical trial paradigm teaches, here, is that to the extent that in etiologic research the concern is, as it usually is, merely with a cause–effect relation of etiologic interest, the particularistic outlook should give way to the abstract one. On the other hand, actual, ultimate questions of etiology, having to do with causal explanation of cases that actually have occurred [1, p. 326], involve an added component issue, namely realization or distribution of the determinant and of modifiers of its effect—a particularistic issue [1, pp. 255–256]. Thus, whatever may be the cause–effect relation between smoking and the incidence/risk of lung cancer (an abstract issue), the etiologic role of smoking is nil in cases, or in times and places, characterized by absence of smoking (a particularistic characterization of the realization or distribution of the determinant).

When the concern is with an etiologically relevant cause–effect relation (abstract) rather than an actual etiologic problem (particularistic), adoption of the clinical trial paradigm as to the irrelevance of time and place leads to major benefits in terms of both validity and efficiency of research, as illustrated in the sections below. These benefits may be viewed as resulting from liberation from the traditional preoccupation with (particularistic) "general populations" or other "target populations" and the associated survey outlook in epidemiologic research. That preoccupation and outlook is natural when epidemiologists are officers of public health and, thus, concerned with community diagnosis about occurrence, epidemic or endemic; but, as has been noted, it is alien to science [7].

In clinical trials, as in all scientific (abstract) occurrence research, there is an expressly designed, particularistic study population, formed within an explicitly selected source population. The study population represents the *domain* (abstract) of the study [1, pp. 44–45] instead of "the general population" or any other particularistic "target population"; is it not construed as a sample of any population [1, Chap. 3]. The clinical scientist's outlook is in accord with that of the laboratory scientist, who does not dream of, or pursue, sampling of the "general" rat population in any community of rats.

DETERMINANT SCALE AND CONTRAST

The traditional epidemiologist, preoccupied with "the general population", commonly studies the effect of a potential etiologic agent by contrasting the exposed segment of that (or a related) "target population" with the remainder in it, that is, by contrasting exposure with nonexposure, with the latter unspecified apart from the inherent absence of exposure [3, Chaps 8–9).

In a clinical trial concerned with the effect of an agent (explanatory trial [4]), treatment with the agent is contrasted with *comparable treatment* without the agent (placebo treatment), not with mere absence of the treatment involving the agent. This is understood to be necessary for isolating, empirically, the effect of the agent from that of the extraneous aspects of the treatment with the agent. Only the *theoretical* contrast is, in its simplest terms, one between presence and absence of the agent [1, p. 30].

The lesson to be gleaned from this, cardinal, feature of the clinical trial paradigm is that, in reference to any loosely defined source population in causality-oriented nonexperimental research, the common two-point scale (exposure vs nonexposure, say) for the determinant in empirical terms is generally indefensible. In the source population there are, at any given moment, people representing the empirical *index* category of the determinant, embodying the agent at issue (cf. treatment with agent); and there are, of course, people not representing the index category—commonly the vast majority.

Of the latter, non-index, segment of the source population, some represent the *reference* category of the empirical scale of the determinant, characterized by *comparability of extraneous effects* with the empirical index category of the determinant [1, pp. 30–31, cf. treatment with placebo]; the remainder—commonly the vast majority—fall in the *extraneous*, "other" category of the determinant's empirical scale (cf. falling outside the realm of the clinical trial).

With this trichotomy pertaining to the source population, the study population must be viewed as a subpopulation of it, consisting of representatives of the index and reference categories but not of the extraneous category, the clinical-trial paradigm being very clear on this. Thus, in proper nonexperimental research on cause–effect relations the study population generally comprises only a small subsegment of the source population (to say nothing about "the general population"—whatever the definition of the latter may be) [1, pp. 29–36, 218–227]. A monument to the still common failure to appreciate the distinction between the non-index range of the empirical scale of the determinant and a proper reference category, as a subsegment of this range, is the concept and problem of "the healthy worker effect"—arising from a contrast between an index population and "the general population" as the reference population [1, pp. 32–33].

That the need to appreciate the essence of the experimental contrast as a paradigm for nonexperimental epidemiologic research is not felt is, perhaps, the prime example of the "double standards" [8] still prevalent between the experimental and nonexperimental modalities of cause–effect research—with major implications for the validity of the latter.

DETERMINANT DISTRIBUTION

Inherent in the traditional contrast-formation in epidemiologic study design is a propensity to allow the distribution of the source ("target") population to determine the distribution of the determinant(s) in the study population-itself [3, Chap. 9]. The famous Framingham Heart Study is an example, and to some perhaps even a paradigm [3, pp. 199–201], of this: the study population was formed as representative sample of the source population, even when selectivity according to various determinants of risk would have been feasible on the basis of the data from the baseline survey. The practice is rooted in the notion that "the general population" or some other "very large population" is an ideal "target" of generalization, and that an ideal study population is derived by probability sampling of such a "target" population [9]. (This outlook, a derivative of the traditional paradigms of epidemiologic research noted above, again reflects defiance, or ignorance, of established principles of science, specifically as to the essence and foundations of scientific generalization [1, pp. 47, 108–109].)

Theory and practice of clinical trials (and of laboratory research) are to the effect that deliberate choice of the distribution (allocation) of experimental units among the compared categories of the determinant is free of any negative influence on validity yet of great relevance for efficiency of study—with equal allocation optimal in the context of any single contrast with equal unit costs between the categories [1, pp. 55–60].

That no allocation occurs in nonexperimental research (by definition) is not a rational basis for not following the clinical trial paradigm. Determinant-selective admission into the study population, within any source population, remains perfectly feasible in most instances [1, pp. 30, 57], and failure to take advantage of this is inexcusable in the context of expensive follow-up. An example of this error of design, much more startling than the Framingham Heart Study, or even the subsequent Collaborative Perinatal Project [10], is a truly mammoth cohort study just recently initiated in China [11].

BASIC TYPES OF STUDY

The most central piece of traditional dogma in epidemiologic research is that there are, in broadest terms, two fundamental types of epidemiologic study (with individuals as the units of observation): the "cohort study", contrasting persons in different categories of a determinant as to subsequent occurrence of (categories of) an outcome phenomenon, and the "case–control study" contrasting persons in different categories of an outcome phenomenon as to an antecedent distribution of a determinant [3, Chaps 8–9].

In an attempt to understand this, consider the clinical trial: the concern is, always, to contrast categories of treatment as to subsequent occurrence of (categories of) some outcome phenomenon, whereas comparing patients in

different categories of outcome as to the antecedent distribution of treatment is uninteresting if not downright perverse [2, 12].

That the "case–control" outlook, as defined, has no place in clinical trials suggests that it has no place in cause–effect research in general. And indeed it does not: in *all* research (applied) on the causation of a phenomenon, a sufficient, and indeed the only proper, concern is to compare the occurrence of categories of this outcome phenomenon among categories of the determinant at issue [1, 2, 12, 13]. To this end one must *always* aim at classifying all members of an explicit study population as to the outcome phenomenon and, in the context of a binary outcome, *always* classify the identified cases according to their histories of the determinant—to obtain the numerators of the empirical outcome rates for the compared categories of the determinant. And with equal force, one must *always* assess the sizes of the corresponding rate denominators in the entire study population, that is, classify its members (over the follow-up) according to the determinant, analogously to the classification of the cases. The latter may be done in absolute (census) or relative (sampling) terms; but either way, the proper concern is to contrast the determinant categories as to outcome rates themselves (involving census denominators) or the corresponding quasi-rates (involving sample denominators); the concern is not to contrast the determinant distributions of the cases and the study population, nor of cases and noncases.

In short, the clinical trial paradigm shows that the notion of a "case–control" study, as defined, is but a fallacy ("trohoc fallacy" [12]) and an illusion [13], and not a true alternative to anything. For the employment of a cohort (closed population) as the study population the true alternative is that of using a dynamic (open) population [1, pp. 48–53]; and, as was observed above, for the census approach to the assessment of rate denominators (in addition to numerators), the alternative is sampling of the study population [12; 1, pp. 69–73]—the census-sample, case–base or case–referent strategy.

Integral to the "case–control" fallacy, to its illusion of a fundamental symmetry in design options, is the traditional notion that the case series and the "control group" of noncases should have identical distributions according to various extraneous determinants of the outcome occurrence or, if they do not, special measures should be taken to avoid bias [3, pp. 167–168]. That this notion is wrong becomes apparent upon examination of the clinical trial paradigm. In a trial with identical age-distributions of subjects between the index (agent) and reference (placebo) treatments, comparison of the index and reference rates without regard for age is not biased on account of age, however strongly age may determine the outcome. By the same token, comparison is also valid, without regard for age, between the corresponding quasi-odds of outcome, for which the case series supplies the numerators and a representative sample (in lieu of census) of the noncases provides the denominators; their ratio is a valid estimate of the actual outcome odds ratio contrasting the two treatments. That the case and non-case series have different distributions by age is, even in the context of unconfounded study base, the expected consequence of age being an extraneous determinant of the outcome rate; the difference is not, in such a situation, a manifestation of confounding, and no cause for remedial action [1, pp. 69–70]. Even if the notion is modified so that the pursuit of identity of distributions between cases and noncases is taken to pertain to predictors of determinant status, or to extraneous outcomes related to the determinant, the presumption of symmetry fails again, as examination of the clinical trial paradigm will readily show.

BASIC OUTLOOK

If there is one outstanding context in which the traditional epidemiologist does emulate the practices in clinical trials, it is the context of the most basic outlook in "etiologic" (generally mere cause–effect) research. He dreams of, and pursues, a cohort obtained as a simple random sample of "the target population", classified as of its formation ("baseline", zero time, t_0) according to "exposure" categories and followed into the future so as to detect subsequent cases—with the aim of relating the incidence of their occurrence to the status at the baseline [3, p. 194]. This is exactly what is done in clinical trials, short of any sampling, naturally, even if complexities arise from noncompliance, for example; and yet this is at variance with the proper basic outlook in etiologic research, the traditionalist—and even revisionist [2]—epidemiologic presumption notwithstanding.

The baseline has meaning in clinical trials but not, generally, in nonexperimental cohort studies. Its meaningfulness in clinical trials

derives from the built-in artifice that the determinant status undergoes intentional perturbation at the baseline only: as a result of subject selection, the determinant commonly has at t_0 a history of constancy; and in any case, the intention is that after this point it is artificially *kept constant* for each subject in its assigned category. In nonexperimental cohorts the determinant commonly has a history of inconstancy before t_0, and in any case the intention is not to keep it constant after t_0. In nonexperimental cohort studies it is, therefore, necessary to account, expressly, for the time course of the determinant both before and after the cohort's t_0, attributing no special import to the determinant status as of the cohort's t_0; indeed, even in a clinical trial outcome occurrence is not actually related to determinant *status* at t_0 (which is in transit), but to futuristic determinant *allocation* at t_0.

There is an even larger fallacy in any attempted emulation of clinical trials in etiologic research, in that the ordinary ("parallel") clinical trial is not suited, even in principle, for being a paradigm in this context. Consider a couple of examples. If all people at early age (t_0) were assigned to, and thereafter continually maintained at, different but constant levels of blood alcohol, one would never learn that it is the level at the time of accident outcome (accident or no accident) that alone matters in etiologic explanation of cases; that the history does not. Similarly, if people at large, at suitably early age, were assigned to, and maintained at, different but constant levels of daily physical activity, one would never learn that the level hours before the outcome in terms of myocardial infarction or no myocardial infarction matters in the sense of causation (precipitation); that the levels during an antecedent period matter in the sense of prevention; and that yet earlier levels do not matter at all.

As those examples suggest, the proper outlook in etiologic research, with keen appreciation of the relevance of time, is that of anchoring the scale of time, individually, to the moment of *outcome* classification; of forming intervals of (individual) time backward from that moment; and, then, thinking of, and treating, the realizations of "the" determinant at the various intervals of retrospective time as realizations for interval-specific *separate* determinants, commonly mutually confounding [1, pp. 33–34, 226–227], of incidence *density*. The dream, properly, is that those specific to ranges of determinants retrospective time do not exhibit an unmanageable degree of collinearity. Preoccupation with a cohort's t_0 [2], as was noted above, is misguided [13] in this context; and the added point was that historical constancy of the determinant as of the time of outcome—inherent in parallel clinical trials—is tantamount to intractable confounding among the separate, time-specific historical determinants of current incidence density.

This inherently retrospective, inherently multi-determinant outlook in meaningful etiologic research even in referrence to a single agent, outlook in which time is inherently anchored to that of outcome classification, cannot be captured by the use of the (parallel) clinical trial as a paradigm.

Acknowledgements—This work was supported by the EMGO Institute, Free University, Amsterdam. Critique of the manuscript by Dr J. Fleiss is gratefully acknowledged.

REFERENCES

1. Miettinen OS. **Theoretical Epidemiology. Principles of Occurrence Research in Medicine**. New York: Wiley; 1985.
2. Horwitz RI. The experimental paradigm and observational studies of cause–effect relationships in clinical medicine. **J Chron Dis** 1987; 40: 91–99.
3. Lilienfeld AM. **Foundations of Epidemiology**. New York: Oxford University Press; 1976.
4. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. **J Chron Dis** 1967; 20: 637–648.
5. Van Laer PH. **Philosophy of Science**. New York: The Ad Press Ltd; 1956: Part I, pp. 72–73.
6. Thompson JA. **Introduction to Science**. Cambridge, Mass.: The University Press; 1911: 40.
7. Friend JW, Feibleman J. **What Science Really Means**. London: George Allen Unwin; 1937: 110–111, 149, 179.
8. Feinstein AR, Horwitz RI. Double standards, scientific methods, and epidemiologic research. **N Engl J Med** 1982; 307: 1611–1617.
9. Moore FE. Committee on design and analysis of studies. **Am J Publ Health** 1960; 50: 10–19.
10. Heinonen OP, Slone D, Shapiro S. **Birth Defects and Drugs**. Littleton: Publishing Sciences Group; 1977: Chap 2.
11. Hammond EC, You W-c, Wang L-d. Possible contribution from epidemiologic studies. **Environ Health Persp** 1983; 48: 107–111.
12. Miettinen OS. The "case–control" study: valid selection of subjects. **J Chron Dis** 1985; 38: 543–549.
13. Miettinen OS. Striving to deconfound the fundamentals of epidemiologic study design. **J Clin Epidemiol** 1988; 41: 709–713.
14. Miettinen OS, Caro JJ. Principles of nonexperimental assessment of excess risk, with special reference to adverse drug reactions. **J Clin Epidemiol** 1989; 42: 325–331