The Clinician's Illusion

Patricia Cohen, PhD, Jacob Cohen, PhD

• There are several diseases, including schizophrenia, alcoholism, and opiate addiction, for which the long-term prognosis is subject to disagreement between clinicians and researchers and also among researchers. Part of this disagreement may be attributable to a difference in the populations they sample. The clinician samples the population currently suffering from the disease (a "prevalence" or census sample), while research samples tend to more nearly represent the population ever contracting the disease (an "incidence" sample). The clinician's sample is biased toward cases of long duration, since the probability that a case will appear in a prevalence sample is proportional to its duration, hence "the clinician's illusion." The statistical mechanism of this bias is illustrated and its consequences detailed. Other sources of sampling bias in clinical and research samples are briefly described and partial remedies are suggested.

(Arch Gen Psychiatry 1984;41:1178-1182)

There are several diseases for which there exists a great disparity in the prognostic expectations of practicing clinicians and investigators of the natural history of the disease. This difference in perspective sometimes leads clinicians to reject formal investigations as invalid or at best irrelevant to the patients they treat. It similarly leads some researchers to view clinicians as unable to surrender clearly invalid ideas about the nature of the problems and prognoses of their patients. Although doubtless some of these discrepancies arise from the difference in their professional roles vis-à-vis patients, another possibly major source of this difference is a function of a sampling problem that has not been generally appreciated.

In this article, illnesses in which this bias is most likely to be found are identified and a neglected methodologic problem that could account for the disparity wholly or in large measure is illustrated. In addition, several other sources of bias or discrepancy that may further complicate and obscure the problem are discussed. Finally, this same problem is shown to be reflected in inconsistencies among research findings. Current research reports in psychiatry often do not allow a proper consideration and correction of the bias. Recommendations for future research and research reports are made.

A 1982 example of a disputed prognosis is available for alcoholism.[1] Although it is not the intent of this article to draw substantive conclusions, it seems likely that the debate regarding the long-term outcome of treatment programs for alcoholics[2-4] has been fueled by the conviction of many clinicians that the research findings of the studies in question did not coincide with their clinical experience. Nevertheless, the preponderance of the research findings with regard to the long-term outcome of alcoholism[5-9] seems to support the Sobells' findings[7] in a general way. That is, they indicate that at least half of all serious alcohol abusers recover completely, many without formal treatment. Furthermore, in contrast to conventional clinical wisdom, a fair proportion of these once-alcoholics have returned to social drinking rather than to abstinence.

A similar debate exists with regard to prognosis in schizophrenia. The traditional view is that schizophrenia is a chronic, deteriorating, and more or less incurable disease. Among clinicians, the most favorable outcome estimate is the "⅓, ⅓, ⅓" rule: one third will get better, one third will stay the same, and one third will get worse. The DSM-III of the American Psychiatric Association is even less sanguine, stating that "a complete return to premorbid functioning is unusual—so rare, in fact, that some clinicians would question the diagnosis. . . ."[10] Thus, the issue of prognosis is addressed, in part, by defining the illness in such a way as to resolve it. In sharp contrast, studies of the natural history of schizophrenic illness suggest a much more favorable prognosis than the "one-third will get better" rule.[11-15]

There are a number of other conditions in which a parallel discrepancy exists in the prevailing view of clinicians and results from clinical trials when compared with the evidence produced by research into the natural history of the condition. For example, Schacter[16] has found what many lay people know firsthand, namely, that many people are obese at some time in their life but return, with or without professional help, to a weight within or close to published normal standards on a more or less permanent basis. Similarly, in contrast to the experience of services designed to help people quit smoking, it has been found that more than 60% of once heavy smokers who tried to quit no longer smoked.[16]

Perhaps even more striking is the view of clinicians, widely shared by the public, of opiate addiction as an incurable state for most, if not all, users. This view was forcefully contradicted by Robins and associates,[17,18] who found that of a sample of Vietnam veterans who were addicted to heroin when interviewed after their return to the United States, 71% were drug free 2½ years later, often without great effort. Of all those who became addicted in Vietnam, even a larger proportion, 88%, avoided relapse over the three years following their return.

It is not our intent to come to any substantive conclusions in this article, however, and readers who disagree with our characterization of findings in any given area are requested to consider the methodologic points being made and their relevance to research in areas familiar to them. Disparities similar to those discussed previously may be expected in investigations of recidivism of juvenile delinquents and adult criminals as well. They may also be found, although probably to a lesser degree, in certain more strictly physical illnesses, such as chronic pain conditions, arthritis, and some cancers, and possibly in many other diseases.

Are clinicians such poor observers? Of course not. They are simply reporting a reasonably accurate abstraction from their experience in the treatment of these conditions. Their day-to-day experience consists of encounters with a group of patients who seek their help. Within this experience are two sources of bias, the census nature of the group and the fact that these cases are necessarily drawn from that subset of patients who have sought help from a particular service provider. The biases attributable to the help-seeking behavior will be discussed later in this article and have been alluded to or discussed in many other places.[19,20] However, the bias due to the census nature of the sample is likely to be less familiar and of even greater consequence in a number of settings, and is therefore the major component of what we call the "clinician's illusion."

The diseases most likely to be subject to the phenomenon of large discrepancies between clinicians' views of the disease and researchers' findings are characterized by great variability in the duration of the illness. Other characteristics (which serve to make long duration possible) are that they are not immediately life threatening and that treatment is either confined to symptom suppression or, if curative in intent, is often not successful.

To illustrate how the variation in duration of the illness may affect the clinician's view, look at a hypothetical disease for which the treatment provides only symptomatic relief. To make the example even simpler, without any loss of generality, let us assume that all persons will be treated for the entire duration of the illness. Suppose further that the distribution of duration of illness in the population of all persons who ever contract it is as displayed in Fig 1.

As can be seen, the distribution posited in Fig 1 is skewed, with a median duration for the entire population of less than six months and one tenth of all cases lasting more than two years. Although this is likely to be a reasonable distribution shape for some of the illnesses discussed, a more symmetrical distribution may be characteristic of others. However, the clinician's illusion is not dependent on this shape but only on the degree of the variability of duration relative to the average duration (literally, on the coefficient of relative variation).

Imagine a clinician viewing persons with one of these illnesses as they appear in the current case load, say, on ward rounds on a given day. The probability of encountering a given patient on this day will be proportional to the duration of this patient's illness. A patient whose illness lasts a year is far more likely to appear on any given day than one whose illness lasts only two months, in fact, six times more likely; and a patient whose illness lasts for ten years is ten times more likely to appear than one whose illness lasts only one year. For this example, each of the few patients making up the 2% of the population of longest duration (X = 128) is 64 times more likely to appear than each of the 40% of shortest duration (X = 2), 32 times more likely to appear than each of the 25% of next shortest duration (X = 4), and so on.

The outcome of this duration-dependent differential probability of being encountered on rounds is that the distribution of illness duration for the rounds sample is markedly different from that of the population made up of people who ever contract the illness. By weighting the percentages in each duration interval in Fig 1 by its mean and recomputing the percentages for the weighted distribution, the duration distribution for the rounds sample of Fig 2 is produced. Perhaps most startling is the outcome of this differential weighting for the typical current clinic group; nearly a quarter of the group is made up of the 2% of the population who have the longest duration, ie, the worst
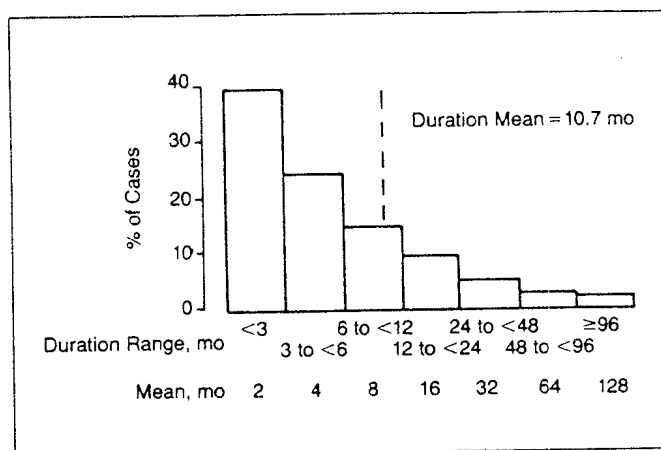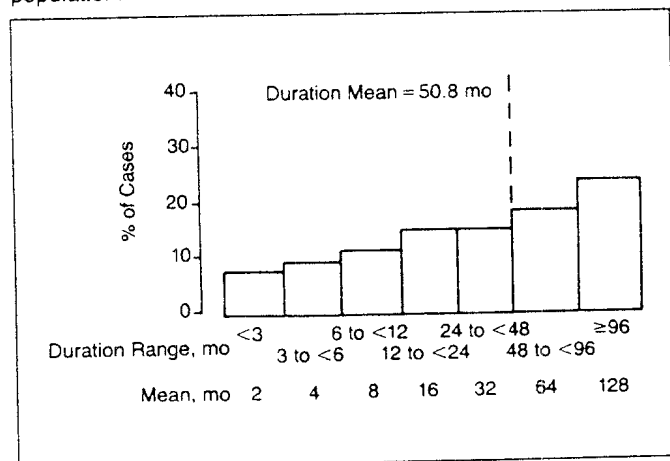

Fig 1.—Distribution of illness duration in population. (Sum of X times percent values for seven intervals in Fig 1 is 1,068. This is now divisor for weighted percentages. For example, percent of lowest interval for rounds sample is 2[40]/1,068 = 7.5%; for next highest, 4[25]/1,068 = 9.4%; etc.)

Fig 2.—Distribution of illness duration in clinic sample from same population.

prognosis. Furthermore, the ultimate average duration for the daily clinic rounds sample seen by our hypothetical clinician will be more than four years (ie, 50.8 months), in stark contrast to the true population average duration of less than one year (10.7 months). Small wonder that the clinician will view any truly population-based findings as grossly discrepant from daily experience. Note again that the phenomenon does not depend on the shape of the distribution but only on its relative variability.

It is useful to examine what happens to this illusion if a duration requirement is inserted into the diagnostic criteria, as has been done for some of the diseases discussed previously. Suppose in our previously mentioned example that those ill less than six months were excluded from the diagnosis. For the remaining 35% of the original population, the mean duration is now 25.4 months. A clinic sample of cases so diagnosed would have a mean duration of 60.5 months, still a large discrepancy.

How then can one correct the observed rounds sample for this distortion of its population representation? Since each case's chance of being included in the sample is proportional to its illness duration, one needs to weight each case inversely by its duration to reproduce the population distribution. Although we do not know what the eventual duration of each case will be, we do know the duration up to the time of observation, and Freeman and Hutchison[21] provide a detailed exposition of how to use durations to date of the clinic sample to estimate the distribution of illness duration for the entire population.

In the previously mentioned example, we assumed a constant probability of treatment throughout the illness. Should this not be the case, the bias may be increased or decreased. If treatment is most likely for the short duration cases, the bias will be lessened, if most likely for the long duration, it will increase. Probably the latter is most frequent in real clinical populations. Again, it is a question of the ratio of probabilities. In our example, if the longest duration group were twice as likely to be treated relative to the shortest, they each would be 128 (rather than 64) times more likely to appear in the sample than a member of the shortest duration group.

One necessary consequence of this bias in the representation in the sample is a bias with regard to the presence of any condition or symptom more characteristic of long-term than short-term treatment of patients. Therefore, it may be impossible to distinguish between a characteristic affecting the duration of illness (failure to recover) from one causally related to the illness.

In sum, the clinician's illusion is the attribution of the characteristics and course of those patients who are currently ill to the entire population contracting the illness. To put the issue in terms familiar to epidemiologists, it is the consequence of using a *prevalence* sample as a substitute for an *incidence* sample.

OTHER SOURCES OF BIAS IN SAMPLES FROM CLINICAL SETTINGS

There are, of course, several other reasons why samples of cases in treatment settings are likely not to be representative of the entire population of persons experiencing the condition in question. Although they have been discussed in other places, it is useful to consider them herein because they are likely to add to the illusion identified previously.

First, those who recover from a given episode of the condition without entering the formal treatment system, that is, by using personal and social resources, are likely to have a better prognosis by virtue of the stability of the curative factors. Unlike those who seek professional help, their help will not disappear when they recover, but will tend to remain through the improved state. A robust constitution is more reliable than the pharmacopeia.

Those who enter the formal treatment system are also likely to have been ill longer. For some of the conditions we have discussed herein, the patient's "social margin" will have declined over time and be accompanied by a decline in the natural support system and increase in alienation from and rejection by others.[22] For these and other reasons, *past* duration of illness may be causally related to *future* duration of illness.

The clinician's illusion is also related to another common observation, at least in the field of schizophrenia and probably in some or all of the other diseases mentioned. This is the apparent association between disease onset at a young age and poor prognosis. To illustrate how this observation could be an artifact closely linked to the previously identified bias, let us examine the characteristics of any given age group seen by our hypothetical clinician, say, all 30-year-old schizophrenics. Note that those who became ill at age 20 years have already had a longer illness duration than those with later onsets, and therefore they can be expected, on the average, to have a duration that projects longer into the future. Or to look at the same phenomenon from another angle, we can note that those who recovered during the decade following their onset at age 20 years (the patients with early onset and good prognosis) are no longer present to be observed. Therefore, if there is any prognostic value to duration of past illness as such, this (remaining) group is, on those grounds alone, more vulnerable to future poor outcome. Note that this is in the absence of any true biologic or any other implication of age of onset for the disease process. However, once again, let us insist that we are not trying to present or evaluate empirical evidence on any issue but only to illustrate problems that arise with empirical findings that are incautiously interpreted.

To return to other reasons why a clinical population might be atypical compared with the entire population, we note another biasing artifact known as Berkson's[23,24] fallacy. It is generally the case that those who have other disabilities that are not causally connected to the condition being investigated are more likely to enter the formal treatment system. Thus, a clinical group may show a conjunction of problems or conditions that are uncorrelated in the population. To use Berkson's example, those who have diabetes *and* cholecystitis are more likely to be found in the hospital than those with diabetes alone. Similarly, patients who are depressed and also abuse alcohol may be more likely to enter into treatment than those with either affliction alone. This error source is a kind of ascertainment bias that, like the others mentioned herein, will further confuse the picture and confound sound conclusions.

This problem should not be confused with that of the distortion of relationships between characteristics that is typical when samples are selected on the basis of these characteristics.[25] When, for example, two characteristics are positively correlated in the population and selection into a sample is based only on the presence of either (or both) of these characteristics, their correlation in the sample will be negative. If other, unrelated, characteristics are also involved in selection, the sign of the correlation may not change, but its magnitude will be diminished. Thus, selection for symptoms is yet another source of discrepancy between clinical samples and the populations whence they come.

FACTORS RELATING TO THE MAGNITUDE OF THE CLINICIAN'S ILLUSION

In addition to characteristics of patients and treatment selection that lead to bias in treated samples, it is worthwhile to identify factors that are likely to maximize the impact of the clinician's illusion. The disparity between what the clinician actually observes and the true state of the population will be greatest when the following occur:

1. The variance in illness duration is large relative to the mean duration.

2. Disease manifestations fluctuate, symptoms are episodic, and treatment is sporadic and handled by multiple clinical services or clinicians. Under these conditions, it is impossible for clinicians to know whether a "cure" is permanent and whether the patient is receiving help in a new setting. Since recovered patients do not return to the clinic whereas relapsed patients do, it may be unclear whether recovery occurs at all.[26]

Clinician's Illusion—Cohen & Cohen

clinical populations. The advisability of studying new cases rather than all current cases should be given careful consideration (but note the risk of accrual shortfall). When the course of the illness is at issue, the prior illness duration must be taken into account. If the past duration is long relative to the follow-up period for any of the patients, an appropriate consideration of the possible effects of this variation on the study conclusions is clearly required.

Second, natural history investigations of cases unselected for duration of illness (and preferably also regardless of treatment) should be considered crucial to the understanding of any illness whose duration varies materially relative to its average length. Such studies are demanding of continuity of effort and support over an extended period of time.

Finally, given the likely distortion of our current knowledge base, we should be cautious in advising lay people in matters relevant to the natural history of chronic diseases. This issue arises with particular potency in the area of genetic counseling. To illustrate, in a research carried out in one of the most sophisticated and research-oriented centers in the country and reported in a journal devoted to research in this area, a parent of a schizophrenic patient is quoted as saying "I've known people who were cured of cancer, of all these illnesses, but I know of no one who was cured of schizophrenia."[35] One would hope that someone concerned with the care of this patient had informed the parent of the high spontaneous remission rate generally found over a more protracted period of time. Nor was any lifetime perspective apparent in the report of the genetic counseling offered. Wider appreciation of the phenomenon we call the clinician's illusion should prove salutary for both clinical practice and the interpretation of research findings.

References

1. Pendery ML, Maltzman IM, West LJ: Controlled drinking by alcoholics? New findings and a reevaluation of a major affirmative study. Science 1982;217:169-175.

2. Marlatt GA: Moderation is the key. Contemp Psychol 1983;28:376-377.

3. Barlow DH, Bellack AS, Buchwald AM, Garfield SL, Hartman DP, Herman CP, Hersen M, Miller PM, Rachman S, Wolpe J: Alcoholism studies. Science 1983;220:554.

4. Armor DJ, Polich JM, Stanbul HB: Alcoholism and Treatment. New York, John Wiley & Sons Inc, 1978.

5. Gottheil E, Thornton CC, Skoloda TE, Waxman H: Etiological Aspects of Alcohol and Drug Abuse. Springfield, Ill, Charles C Thomas Publisher, 1983.

6. Heather N, Robertson I: Controlled Drinking. London, Methuen Publishers Ltd, 1981.

7. Pattison E, Sobell M, Sobell L: Emerging Concepts of Alcohol Dependence. New York, Springer Publishing Co Inc, 1977.

8. Gottheil E, Thornton CC, Skoloda TE, Alterman AI: Follow-up of abstinent and nonabstinent alcoholics. Am J Psychiatry 1982;139:560-565.

9. Vaillant GE, Milofsky ES: Natural history of male alcoholism. Arch Gen Psychiatry 1982;39:127-133.

10. American Psychiatric Association, Committee on Nomenclature and Statistics: Diagnostic and Statistical Manual of Mental Disorders, ed 3. Washington, DC, American Psychiatric Association, 1980.

11. Bland RC, Orn H: 14-year outcome in early schizophrenia. Acta Psychiatr Scand 1978;58:327-338.

12. Bleuler M: The long-term course of the schizophrenic psychoses. Psychol Med 1974;4:244-254.

13. Ciompi L: Catamnestic long-term study on the course of life and aging of schizophrenics. Schizophr Bull 1980;6:606-618.

14. Engelhardt DM, Rosen B, Feldman J, Engelhardt JZ, Cohen P: A 15-year followup of 646 schizophrenic outpatients. Schizophr Bull 1982;8:493-503.

15. Huber G, Gross G, Schuttler R: Long-term follow-up study of schizophrenia: Psychiatric course of the illness and outcome. Acta Psychiatr Scand 1975;52:49-57.

16. Schacter S: Recidivism and self-cure of smoking and obesity. Am Psychol 1982;37:436-444.

17. Robins L: The Vietnam Drug User Returns. Special Action Office Monograph, May 1974 (series A, No. 2).

18. Robins L, Helzer JE, Hesselbrock M, Wish E: Vietnam veterans three years after Vietnam, in Brill L, Winick C (eds): The Yearbook of Substance Use and Abuse. New York, Human Sciences Press, 1980, vol II.

19. Kleinbaum DG, Kupper LL, Morgenstern H: Epidemiologic Research: Principles and Quantitative Methods. Belmont, Calif, Wadsworth Publishing Co, 1982.

20. Susser M: Causal Thinking in the Health Sciences. New York, Oxford University Press, 1973.

21. Freeman J, Hutchison GB: Prevalence, incidence and duration. Am J Epidemiol 1980;112:707-723.

22. Gruenberg EM: The social breakdown syndrome: Some origins. Am J Psychiatry 1967;123:1481-1489.

23. Berkson J: Limitations of the application of fourfold table analysis to hospital data. Biomet Bull 1946;2:47-53.

24. Berkson J: The statistical study of association between smoking and lung cancer. Proc Staff Meetings Mayo Clinic 1955:319-348.

25. Dawes R: A case study of graduate admissions: Application of three principles of human decision making. Am Psychol 1971;26:180-188.

26. Vaillant GE: The distinction between prognosis and diagnosis in schizophrenia: A discussion of Manfred Bleuler's paper, in Wynne LC, Cromwell R, Matthysse S (eds): The Nature of Schizophrenia. New York, John Wiley & Sons Inc, 1978, pp 637-639.

27. Kraemer HC: Coping strategies in psychiatric research. J Consult Clin Psychol 1981;49:309-319.

28. Fletcher RH, Fletcher SW: Clinical research in general medical journals: A 30-year perspective. N Engl J Med 1979;301:180-183.

29. Gift TE, Strauss JS, Harder DW, Kokes RF, Ritzler BA: Established chronicity of psychotic symptoms in first admission schizophrenic patients. Am J Psychiatry 1981;138:779-784.

30. Leske MC, Ederer F, Podgor M: Estimating incidence from age-specific prevalence in glaucoma. Am J Epidemiol 1981;113:606-613.

31. Louis TA, Shapiro SH: Critical issues in the conduct and interpretation of clinical trials. Annu Rev Public Health 1983;4:25-46.

32. Mednick SA: Berkson's fallacy and high-risk research, in Wynne L, Cromwell R, Matthysse S (eds): The Nature of Schizophrenia. New York, John Wiley & Sons Inc, 1978, pp 442-452.

33. Kahneman D, Slovic P, Tversky A: Judgement Under Uncertainty: Heuristics and Biases. Cambridge, England, Cambridge University Press, 1982.

34. Kupfer DJ, Rush AJ: Recommendations for depression publications. Arch Gen Psychiatry 1983;40:1031.

35. Schultz PM, Schultz SC, Dibble E, Targum SD, vanKamen DP, Gershon ES: Patient and family attitudes about schizophrenia: Implications for genetic counseling. Schizophr Bull 1982:8:504-513.