HEROIN ADDICT FORECASTING-STATISTICAL FALLACIES AND DISPUTED FACTS

Alex Richman and Vincent V. Richman
Department of Psychiatry, Beth Israel Medical Center and the Mount Sinai School of Medicine of the City University of New York.

"There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact."

Twain, M.

INTRODUCTION

Heroin addiction is a major problem. The data base for assessing the true dimensions of the heroin problem is, at best, woefully inadequate, and, at worst, a horrendous scandal. (Senate Committee on Governmental Operations, 1973). This paper is specifically concerned with methods currently used for assessing and forecasting extent of heroin addiction.

"PEAK YEAR" OF ONSET OF HEROIN ADDICTION

In 1972 the Special Action Office for Drug Abuse Prevention (SAODAP) began to contend that the rate of growth of new heroin addiction had been "slowing"since 1969. "Asked what the addict population rose from in 1965 and what it peaked at in 1969, a SAODAP spokesman said they had no absolute numbers to go on, only percentage estimates." (Drugs and Drug Abuse Education Newsletter, 1972).

Dupont and Greene (1973) analyzed year of first heroin use for admissions between July 1970 to December 1972 to Narcotics Treatment Administration, District of Columbia. Dupont and Greene tested the hypothesis that a lag between onset of heroin use and entry into treatment explained the drop in "addiction incidence" by graphing the distribution of year of first use for successive cohorts of admissions over five consecutive six month periods and obtaining curves with peaks at 1969 for the latest four subgroups. "If the hypothesis were correct, one would expect the peak of these five curves to shift over time...this suggests that the decline in incidence is real and not related to delay in seeking treatment."

STATISTICAL PROBLEMS IN ASSESSING "PEAK YEAR"

1. Changes in incidence cannot be inferred from first admission data. Changes in first admission rates do not prove that incidence has changed. Kramer (1957) has shown that changes in first admission rates reflect changes in incidence ONLY when the following ratio is known to be the same in both time periods.

$$\frac{\text{TOTAL NUMBER OF FIRST ADMISSIONS TO TREATMENT FACILITIES}}{\text{NUMBER OF NEW CASES IN THE COMMUNITY}}$$

Calculation of this ratio requires incidence data from the community, which if available, would obviate any need for using substitute sources. In other words, first admission data reliably reflect incidence ONLY when incidence is known and the above ratio calculated.

2. Misleading effect of use of percentages. The number of admissions of long-term addicts varies from time to time, and will thus affect the percentage with shorter durations of addiction (Richman, 1974). On the other hand, fluctuation in the total number results in similar percentages representing quite different numbers.

3. Duration of addiction and incidence. A decrease in the absolute number of admissions of recent onset might be due to a decreased percentage of new cases entering treatment or a delay in seeking treatment or the effect of programs' admission policies or a decrease in incidence in the community.

WHAT TO "PEAKS" REPRESENT?

"Peaks" in the distribution of year of onset of heroin use among admissions have been attributed by Dupont and Greene (1973), Jaffe (1973) to peaks in incidence. These peaks can be explained from statistical theory.

Natural, economic and sociological data based on measurements show a distinct tendency to group about a given point. This grouping tendency gives a rise to "peak" which always occurs in frequency distributions, Arkin and Colton, (1958). It is believed that 80 per cent of all variables data can be at least approximately represented by the normal distribution function. This distribution occurs when four or more input variables operate randomly and independently and whose effects combine to give an out-put, King (1971).

Jaffe (1973) showed "New York" data with "peaks" of heroin use in 1967 and 1968. These "New York" data were derived from two agencies reporting to the Drug Abuse Reporting Program (NIMH and Institute of Behavioral Research, Texas Christian University). Figure A shows the distribution for admissions to each agency during 1970, 1971 and 1972, by year of first illegal narcotic use. There were marked differences in the distribution of year of onset for each agency with the "peaks" varying for each agency from year to year.

The marked "peaks" for "New York" result from the data for two diverse agencies being summed together. The greater the number of random and independent variables which are combined, the more likely the output will result in a normal distribution function with a "peak".

LOG NORMAL DISTRIBUTION OF INTERVALS BEFORE IDENTIFICATION

Incubation periods of infectious disease

The incubation period is a characteristic common to infectious diseases and to altered physiological states produced by chemical and physical agents.

Sartwell (1950) demonstrated that the distribution of incubation periods forms a consistent pattern in a number of human diseases; the usual frequency curve of incubation time taking the form of a logarithmic normal curve, both for diseases with very short and very long incubation periods.

Ascertainment intervals of heroin addiction

Heroin addicts are identified by ways other than admission to a clinical treatment program, e.g., arrest, questionnaire or urinalysis. I suggest that "ascertainment interval" be used to describe the interval between the onset and later identification in order to emphasize that we are dealing with identification and not incidence.

This ascertainment interval would be affected not only by personal characteristics and the clinical features of the disorder, but by nosocomial factors and threshold-affecting factors. Nosocomial factors refer to institutional changes which affect the ability to accept patients (changes in bed capacity, increased personnel, shorter duration of treatment). Threshold affecting factors refer to the factors which affect whether and when an admission occurs; these include social pressures, fear of arrest or legal compulsion; expectations of effective treatment; attitudes to the treatment program. (Svendsen)

Logarithmic normal distributions of ascertainment intervals for heroin addicts

The intervals between onset and admission were plotted on logarithmic normal probability paper, cumulative percentages being plotted against the logarithm of duration of addiction in years, and straight lines fitted by inspection. Log normal distributions of ascertainment intervals are shown in Figure B for admissions to Boston City Hospital 1971, Haight Ashbury Free Medical Clinic 1971, M.J. Bernstein Institute 1965 and 1971, Narcotics Treatment Administration, D.C., 1970-1971 and Jaffe's data for "New York City" 1971. In view of the diverse clinical programs, geographic areas, and time periods, the linearity of all these distributions between the range of 2-5 years is striking.

Admission Threshold and the Log-Normal Distribution

Progressive changes occurred for the admission threshold at the Beth Israel Medical Center where the median duration of addiction changed from about 2 1/2 years in 1969 to four years in 1972. This delay in admission affected each of the onset years; and resulted in the three admission cohorts retaining their log normal distribution of ascertainment intervals.

The log-normal distribution for durations of addiction arises from the central limit statistical theorem, which states that the sums of independent random variables tend to become normally distributed as the number of individual variables summed becomes very large. The log time distribution means the factors are being summed in log time, which is equivalent to multiplication in linear time.

I believe these log normal distributions result from the complex interactions between individuals, society, and treatment institutions which determine when admission occurs - the threshold affecting factors which are reflected in the time interval between onset and entry of the narcotic abuser into treatment. Rather than indicating community changes in incidence, as measures of admission threshold - these distributions indicate the balance between need for treatment and admission, which is affected by a great many factors.

Changes in Admission Threshold in Washington D.C.

The major, rapid decrease in first admissions to the Narcotic Treatment Agency D.C. in 1972 - Despite their claim to the contrary, Dupont and Greene's data demonstrate a lag in entering treatment. The log normal distributions of ascertainment intervals for admission to the Narcotics Treatment Administration are shown in Figure C. In late 1972 there was a marked decrease in the admission of patients of all durations of addiction, not just a decrease in those with onset of heroin use after 1969.

"LAG" BETWEEN ONSET AND ADMISSION

The delay between onset of heroin use and subsequent entry into treatment is referred to as "lag".

Recently "lag" data have been used to make projections of future drug use and treatment for a given program. Hunt 1974, claims that "lag" is stable from time-to-time, consistent from place-to-place and can be estimated from onset cohorts.

Figure D (from Hunt) is based on data from the Drug Dependence Unit, Connecticut Mental Health Center and is said to demonstrate that: "The distribution often appears stable from year to year (i.e., the slope of each curve in this family of curves is similar), suggesting that there has been little change in addicts' disposition to volunteer for treatment..."

With the cooperation of the Drug Dependence Unit, Connecticut Mental Health Center (H. Kleber, M.D.) and the Drug Abuse Report Program, Texas Christian University (S.Sells, Ph.D.), it has been possible to analyze the source data of Figure D and construct Figure E.

Figure E shows that there was a progressively increasing delay for successive admission cohorts. Persons admitted in 1972 showed greater "lag" than those admitted in 1971 and those admitted in 1971 showed more "lag" than those admitted a year earlier.

Thus, the admission threshold for New Haven was not a persistent or stable characteristic upon which to make projections. This instability from year to year, of course, invalidates the next step.

Hunt states that if onset cohort lag curves are stable, percentages of each year's entrants may be averaged to yield a mean cumulative entry curve, which is a more accurate estimate of overall behavior than any single year.

Figure D-2 (from Hunt) is said to be the "average cumulative entry" curve for Figure D-1. Note that the abscissa is similar for both Figures D-1 and D-2 - years from onset to entry into treatment.

While Figure D-1 shows that some patients with onset of heroin use in 1960, 1961, 1962, 1963, 1964 and 1965 were not admitted until six or more years had elapsed; the "average cumulative entry" curve in Figure D-2 shows that 100% had been admitted within 5 years of onset. If Figure D-2 is superimposed on Figure D-1, it is evident that the "average" fits the 1968 onset cohort, the left-most curve and not the curves for any of the earlier onset years.

This discrepancy between Figures D-1 and D-2 results from Hunt confusing duration of addiction of admissions with duration of operation of the clinical program. Hunt's work table (1975) uses the ratio:

$$\frac{\text{NUMBER ADMITTED IN FIRST YEAR OF PROGRAM}}{\text{ALL ADMISSIONS WITH ONSETS BEFORE PROGRAM OPENED,}}$$

as if it were equivalent to the proportion of an onset cohort who would be admitted within the first year of heroin use.

In his work table the number of entrants during 1970 (with onsets ranging from pre-1960 to 1970) is taken as the numerator for calculating the percentage of an onset cohort who would be admitted within one year of onset. This confusion between admission cohorts (proportion of all onsets admitted in first year of program) and onset cohorts (proportion admitted within first year following onset of heroin use) is particularly striking since Hunt recognizes the difference:

> "it has sometimes been argued that cumulative entry can be better estimated from intake cohorts than from onset cohorts, because an intake cohort is complete...The flaw in this method is obvious".

The analyses and methods used by Hunt have many inconsistencies, and assumptions which are not supported by data. Yet the method has been claimed to be one of the best techniques available today for treatment program planning and for assessing the incidence of drug abuse in the community. (Greene, 1974)

DISCUSSION

Shaw's 1972 Presidential address to this Association differentiates:

1) Statistics singular-
   -the mathematics of the collection, organization and interpretation of numerical data...used with a singular verb.
   -concerned mainly with means.

2) Statistics plural-
   -a collection of numerical data...used with a plural verb.
   -concerned with ends as well as means.

3) Statistics political (using political in its broad sense, not in its partisan sense).
   -political science dealing with state affairs.
   -concerned primarily with ends.

In the field of heroin addiction, it seems as if statistics has been concerned mainly with ends, that good methodology has been more honored in the breach than in observance.

Both users and practitioners of statistics should strive to meet Shaw's requirements for responsibility for-

- insisting on quality
- curbing misuse
- adequate description
- stressing limitations
- helping to make "statistics political" a respected tool for analyzing and resolving problems, not just a numbers game.

SUMMARY

Forecasting requires reliable methods and valid data. While estimates of the prevalence of heroin addiction have been based on inadequate data, faulty assumptions, unstated premises and distorted interpretations, trends in incidence have been inferred from two misleading measures: the "peak year" of onset of heroin use and "lag" between first use and entry into treatment.

Both "peak year" and "lag" are derived from the distribution of intervals between onset and admission - the ascertainment interval. The distribution of ascertainment intervals is a function of admission threshold-affecting factors, (not directly related to incidence) as well as, onset and persistence of heroin use and its transition to heroin addiction.

The records of admissions to treatment have been used for forecasting treatment needs; these forecasts are based on the claim that the distribution of ascertainment intervals remains fairly constant over time. The fallacies of this approach are discussed.

Users and practitioners of statistics on heroin addiction should be more concerned with means and less with ends.

REFERENCES

1. Arkin H and Colton RR: Statistical Methods, Barnes & Noble, Inc., New York City, 1958.

2. Drugs and Drug Abuse Education Newsletter, October, 1972.

3. Dupont RL, Greene MH: The dynamics of a heroin addiction epidemic. Science 181: 716-718, 1973.

4. Greene MH et al: An assessment of the Diffusion of Heroin Abuse to Medium-Sized American Cities, Special Action Office Monograph Series A, No. 5, Washington, D.C.: Special Action Office for Drug Abuse Prevention, 1974.

5. Hunt L: Heroin Epidemics - A Quantitative Study of Current Empirical Data, Washington, D.C., Drug Abuse Council.

6. Hunt LG: Drug Incidence Analysis. Special Action Office Monograph Series A, No. 3, August 1974, Washington, D.C., Special Action Office for Drug Abuse Prevention, 1974.

7. Hunt LG: Admission Records. A tool for treatment program planning, Washington, D.C. Drug Abuse Council, 1975.

8. Jaffe J in Harris TG, "As far as Heroin is concerned, the worst is over" a conversation about the drug epidemic with Jerone Jaffe, Psychology Today. 7: 68-79, et passim; August, 1973.

9. King Jr: Probability Charts for Decision Making, New York, Industrial Press, 1971.

10. Kramer M: A discussion of the concepts of incidence and prevalence as related to epidemiological studies of mental disorder, Am J Pub Health 47: 826-836, 1957.

11. Richman A: Is the onset of narcotic addiction among treated patients in the United States really declining? In Conference on the Epidemiology of Drug Abuse, San Juan, Puerto Rico, February 12-14, 1973.

12. Richman A: Epidemiologic assessment of changes in the onset of narcotic addiction, In: National Research Council. Committee on Problems of Drug Dependence: Report of the Thirty-Sixth Annual Scientific Meeting, March 10, 1974, Washington, D.C., 1974, pp. 669-690.

13. Richman A: Peak year of onset of heroin use: Epidemiologic fact or mystic statistic. In: Developments in the field of Drug Abuse - Proceedings of the National Drug Abuse Council, 1974, Senay E, Shoety V, Alksne H, eds. Cambridge, Mass., Schenkman Publishing Company, Inc., 1975, pp. 254-267.

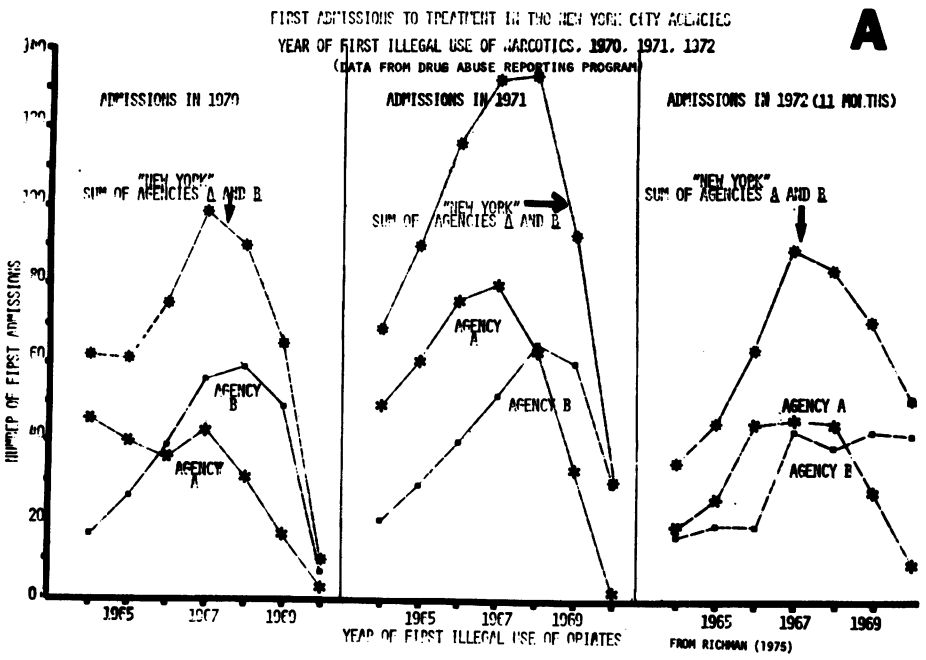
14. Sartwell PE: The distribution of incubation periods of infectious disease, Amer. Jour. Hyg. 51: 310-318, 1950.

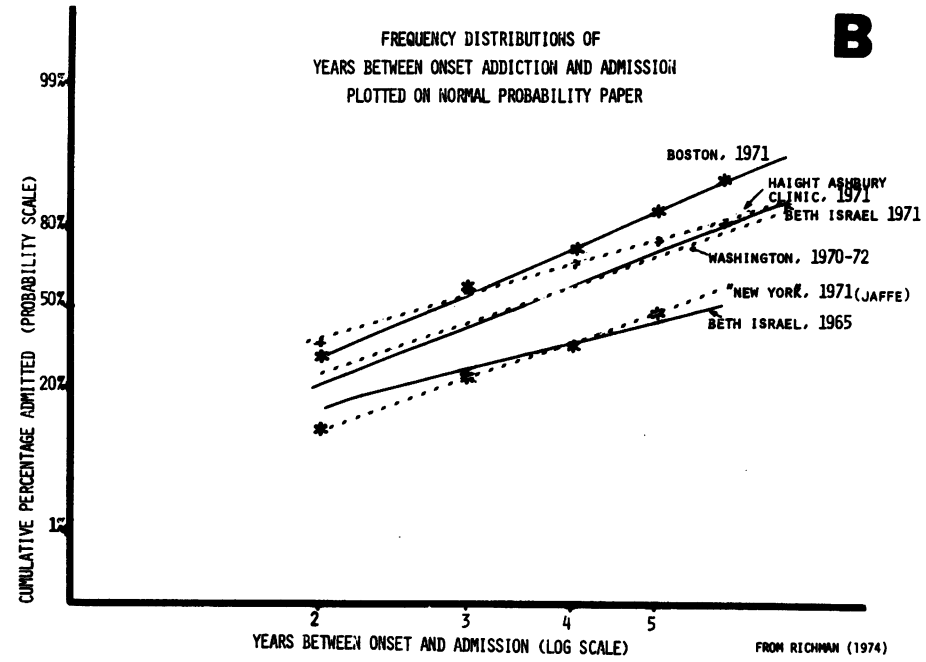
15. Sartwell PE: The incubation period and the dynamics of infectious disease, Amer. Jour. Epid. 83: 204-216, 1966.

16. Shaw WH: Paradoxes, problems and progress. Jl A Stat Assocn. 68: 7-10, 1973.

17. Svendsen B, cited by Stromgren E: Statistical and genetical population studies with psychiatry. Methods and principal results. Congress international de psychiatrie, 6: 155-192, Paris.

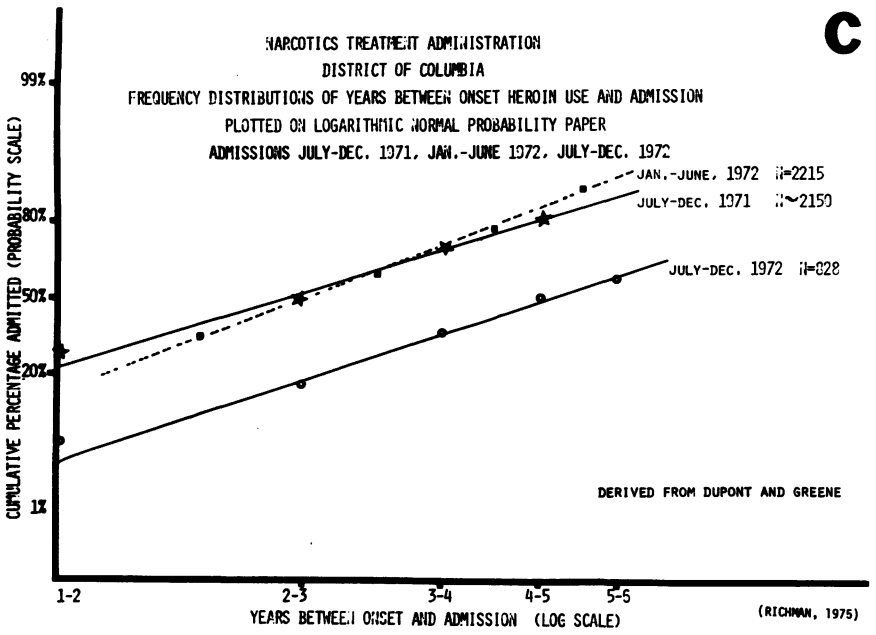
NOTE: FULL SIZE ILLUSTRATIONS AND TEXT CAN BE OBTAINED FROM THE AUTHOR - 307 SECOND AVENUE, NEW YORK, 10003.


A
FIRST ADMISSIONS TO TREATMENT IN TWO NEW YORK CITY AGENCIES
YEAR OF FIRST ILLEGAL USE OF NARCOTICS, 1970, 1971, 1972
(DATA FROM DRUG ABUSE REPORTING PROGRAM)
ADMISSIONS IN 1970
ADMISSIONS IN 1971
ADMISSIONS IN 1972 (11 MONTHS)
"NEW YORK"
SUM OF AGENCIES A AND B
AGENCY A
AGENCY B
NUMBER OF FIRST ADMISSIONS
0
20
40
60
80
1965
1967
1969
YEAR OF FIRST ILLEGAL USE OF OPIATES
FROM RICHMAN (1975)


B
FREQUENCY DISTRIBUTIONS OF
YEARS BETWEEN ONSET ADDICTION AND ADMISSION
PLOTTED ON NORMAL PROBABILITY PAPER
BOSTON, 1971
HAIGHT ASHBURY CLINIC, 1971
BETH ISRAEL 1971
WASHINGTON, 1970-72
"NEW YORK", 1971 (JAFFE)
BETH ISRAEL, 1965
99%
80%
50%
20%
1%
CUMULATIVE PERCENTAGE ADMITTED (PROBABILITY SCALE)
2
3
4
5
YEARS BETWEEN ONSET AND ADMISSION (LOG SCALE)
FROM RICHMAN (1974)

C

NARCOTICS TREATMENT ADMINISTRATION
DISTRICT OF COLUMBIA
FREQUENCY DISTRIBUTIONS OF YEARS BETWEEN ONSET HEROIN USE AND ADMISSION
PLOTTED ON LOGARITHMIC NORMAL PROBABILITY PAPER
ADMISSIONS JULY-DEC. 1971, JAN.-JUNE 1972, JULY-DEC. 1972



(RICHMAN, 1975)

D1

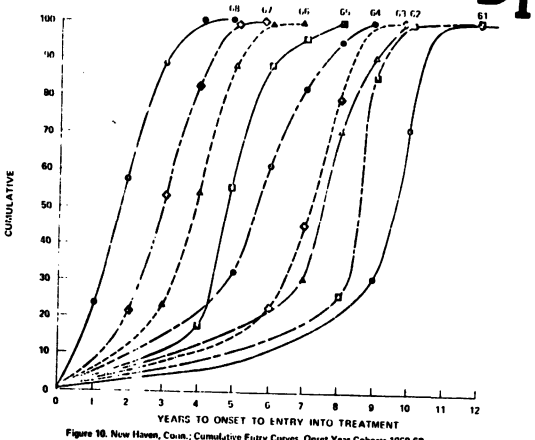


Figure 10. New Haven, Conn.: Cumulative Entry Curves, Onset Year Cohorts 1960-68
(Last Year Datum Incomplete, Data through March, 1973)
From L.G. Hunt, op. cit.

D2

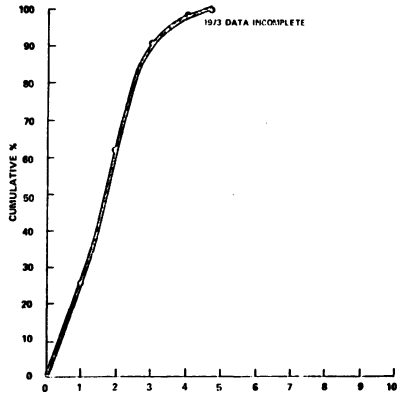


Figure 13. New Haven, Conn. Average Cumulative Entry Curve (Based on Fig. 10)
From L.G. Hunt, op cit.

E

ADMISSIONS TO DRUG DEPENDENCY UNIT
CONNECTICUT MENTAL HEALTH CENTER
(REPORTED TO DRUG ABUSE REPORT PROGRAM)

FREQUENCY DISTRIBUTIONS OF YEARS BETWEEN ONSET OPIATE USE AND ADMISSION
PLOTTED ON LOG NORMAL PROBABILITY PAPER BY YEAR OF ADMISSION