Long-Term Effects of Neuroleptics
(Adv. Biochem. Psychopharmacol., Vol. 24),
edited by F. Cattabeni et al.
Raven Press, New York © 1980.

Use of Survival Curves in Analysis of Antipsychotic Relapse Studies

John M. Davis, Maurice W. Dysken, *Shelby J. Haberman, Javaid I. Javaid, Sidney S. Chang, and Grant A. Killian

*Illinois State Psychiatric Institute, and *University of Chicago, Chicago, Illinois 60612*

The basic clinical concern in the use of maintenance medication is the prevention of illness relapse. To date, there have been 30 maintenance studies that have compared the efficacy of antipsychotic drugs versus placebo in schizophrenic patients. All 30 studies report that patients receiving drugs experience fewer relapses than those maintained on placebo. Since all of these studies are addressing the identical hypothesis, their separate results can be combined. The probability that the common finding could occur by chance is substantially less than 10^{-100}. In addition, these studies consistently report a relapse figure of about 50% for those patients taking placebo. It is stated frequently in psychiatric textbooks and articles that since 50% of the patients do not relapse on placebo, 50% of the patients do not need drugs. Indeed, this statement is granted the force of law, as demonstrated by certain regulations in the State of California. It is this statement that we want to question. The purpose of this chapter is to challenge the view that 50% of patients do not relapse. This conclusion is based on the studies cited above, based on results from an invalid conceptual appraisal of that data. We were the first to raise such considerations of exponential relapse rate, and this chapter will develop the argument more fully than our original paper (2).

Let us start by looking at the data descriptively. Perhaps the most complete study is that of Hogarty and Goldberg (5) who compared maintenance medication against placebo in outpatient schizophrenics. Cumulative relapse rates are provided in their paper and expressed here in Fig. 1 (A) as the percentage of patients unrelapsed. In doing this, we have been guided by such analogies as drug half-life *in vivo* or half-life of radioactive decay. Note that the percentage of patients unrelapsed does not yield a straight line when plotted over time, but rather curves and progressively flattens out. The simplest assumption is that the percentage of patients who relapse per unit time is constant. We have a mathematical problem, however, in that as patients relapse they are dropped from the study so that the number of patients left unrelapsed, who have the potential to relapse, decreases with time. In other words, the number of patients

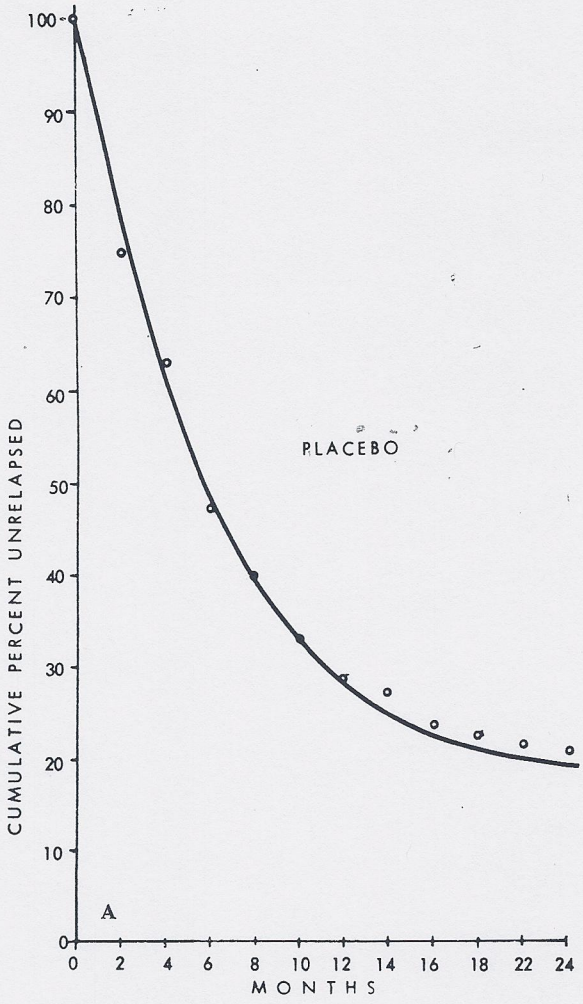
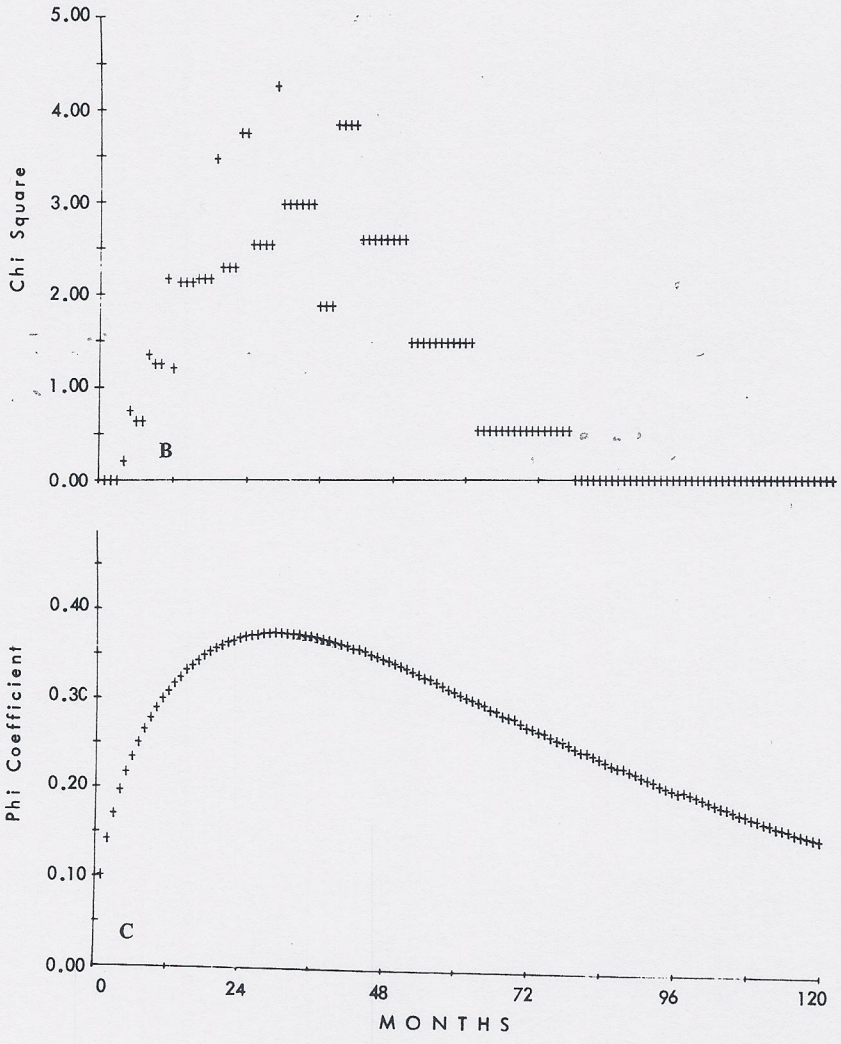


FIG. 1. Cumulative relapse over time is expressed as the % of unrelapsed over time. Data: Hogarty and Goldberg (5). Figure 1B represents the chi square plotted over time with a relapse rate of 0.026 (drug) and 0.071 (placebo) ($N = 30$). This figure compares the projected number of patients who relapsed or did not relapse per month on either drug or placebo at these relapse rates. Those who relapsed or did not relapse are examined monthly in a 4-fold chi square table (with Yates corrections). This shows that if a trial is conducted over an approximate time period of less than 1½ years, the chi square test is insensitive in detecting a drug-placebo difference. Similarly, if the trial lasts over 4 years, the chi square analysis is again insensitive. The chi square is clearly artifactually a function of time.

Figure 1C represents the phi coefficient plotted over time with the relapse rate of 0.026 (drug) and 0.071 (placebo) ($N = 30$). The curve indicates the phi coefficient comparing patients who relapsed or did not relapse while on drug or placebo (1–120 months). The phi coefficient with a large sample size smoothes out the discontinuities. To compute the number of patients who relapse each month, the integer value of the expression e^{-Kt} when K = relapse rate and t = time is calculated. Since the integers are discontinuous, the chi square



with a small sample size is discontinuous, but the phi coefficient with a large sample size is essentially continuous. Both figures indicate that the chi square or phi is a function of time, a particularly inappropriate index of relapses at trials under 1½ years or over 4 years in length.

in the study is shrinking with each increment of time. Let us say that the percentage of patients who relapse is constant over time, for example, 15% per month. As a gross estimate, we might say that out of an initial 100 patients, 15% will relapse (15 patients) at the end of 1 month, leaving a total of 85 patients. In the next month, 15% of the group of 85 patients (13 patients)

will relapse, leaving 72 patients. After the next month, we would be left with 61 patients, 11 patients having relapsed. Note that the absolute number of patients relapsing shrinks. Of course, this goes on every day since the relapse risk of 15% per month is the same as 0.5% per day. Thus we could perform essentially the same calculations starting out with 0.5% of the number of patients relapsed from the study after the first day. This is the same mathematical approach that is involved in compound interest. It is compounded not monthly, not daily, but continuously; hence if the relapse rate is 15% per month, then the number of patients unrelapsed at the end of the first month would be 86, the second month 74, and, similarly, 63, 54, 47, 41, 35, 30, 25, 22, 19, and 17 for the 12 months of the first year.

Data on the number of patients relapsing per unit of time are available in all three of the studies (1,5,7); we have plotted these data as a logarithm of the percentage of patients unrelapsed over time, for all three studies, in Figs. 2, 3, and 4. Note that Prien and Cole (7) present data only on placebo over time and do not present what happens to patients maintained on drugs. The other investigators, Hogarty and Goldberg (5) and Caffey et al. (1), do present the number of patients relapsing on drug and on placebo. It is easy to see that far more patients relapsed on placebo than on drug. To look at the rate of relapse, the fraction of the patient population relapsing per unit time is divided by the number of patients in that study within that period of time (see bottom panels of Figs. 2, 3, and 4). To inspect the cumulative relapse rate, the number of patients unrelapsed is corrected for the decreasing sample size by calculating the natural logarithm of that percentage of patients unrelapsed. Note that in Figs. 2, 3, and 4, all these curves show approximate linearity.

To return to the 30 studies comparing drug and placebo relapses, all of these studies present similar data such that a 2×2 chi square table can be computed from the number of patients unrelapsed and relapsed receiving drug or placebo at the end of the study. The majority of the studies are 4 or 6 months in duration. With the exponential relapse observed in Figs. 2 through 4, one would predict that the number of patients relapsed by 4 to 6 months would be around 50%. Since almost all of the studies use what we call the 2×2 table (chi square) method, we will discuss its limitation using the following hypothetical example. If the relapse rates of the placebo and drug groups are 7.1% and 2.6%, respectively, per month, chi squares can be calculated for each month over a 10-year period. This example is calculated with rates similar to that observed in Fig. 2. For example, chi square at a given point in time is essentially a function of the distance between the two curves (note in Fig. 2 the difference at early points, such as at 3 or 6 months). Compare the difference at 1 or 2 years and again at 3 or 4 years. In Figure 1 (B) we have presented a plot of a chi square with the Yates correction based on $N = 30$. Note that the chi square varies over time. If the study is brief in duration, such as 6 months or 1 year, the results are insignificant; after several years, results are significant. The chi square would be unimpressive when the study was extended for a

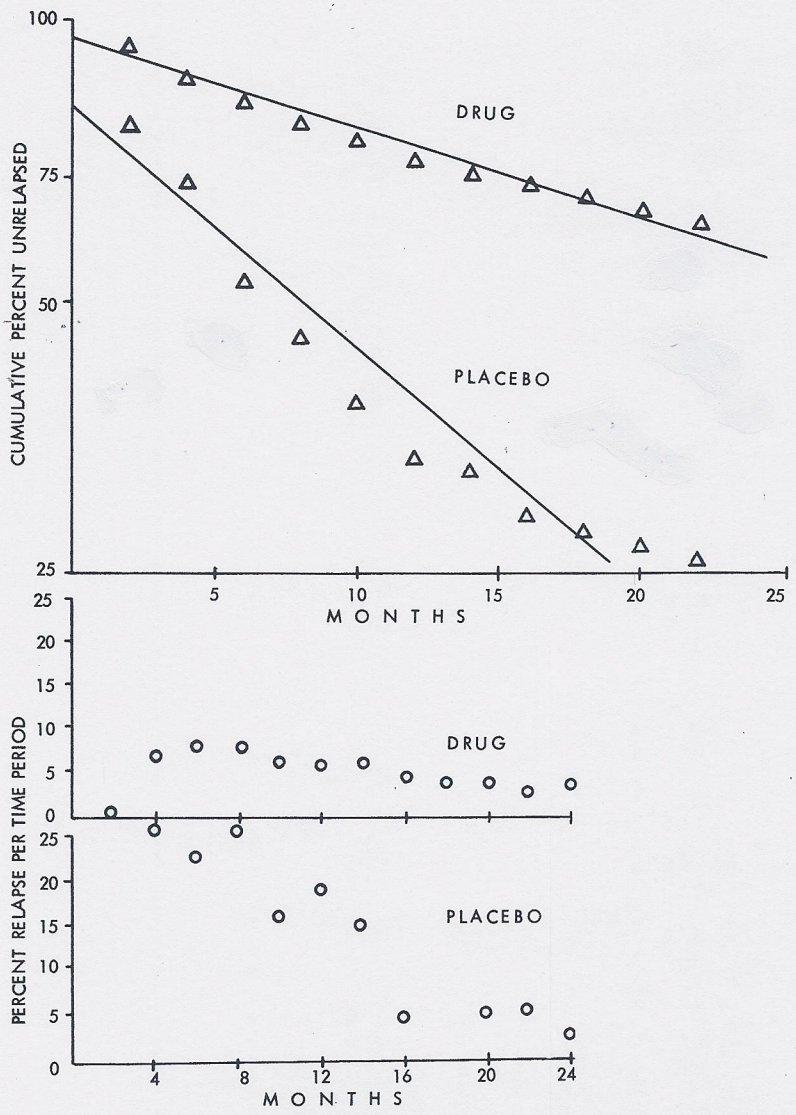


FIG. 2. The survival distribution of cumulative relapse over time is expressed as the Ln_e of % unrelapsed over time. Data: Hogarty and Goldberg (5). Bottom panel: Hazard rates, same data. Number of relapses divided by the number of patients being studied at each time period.

long follow-up period. The reason for this becomes clear upon inspection of Fig. 2. During the first few months of the study, the drug-placebo difference increases. However, toward the end of the study, when essentially almost all the placebo patients have relapsed, the drug patients are continuing to relapse

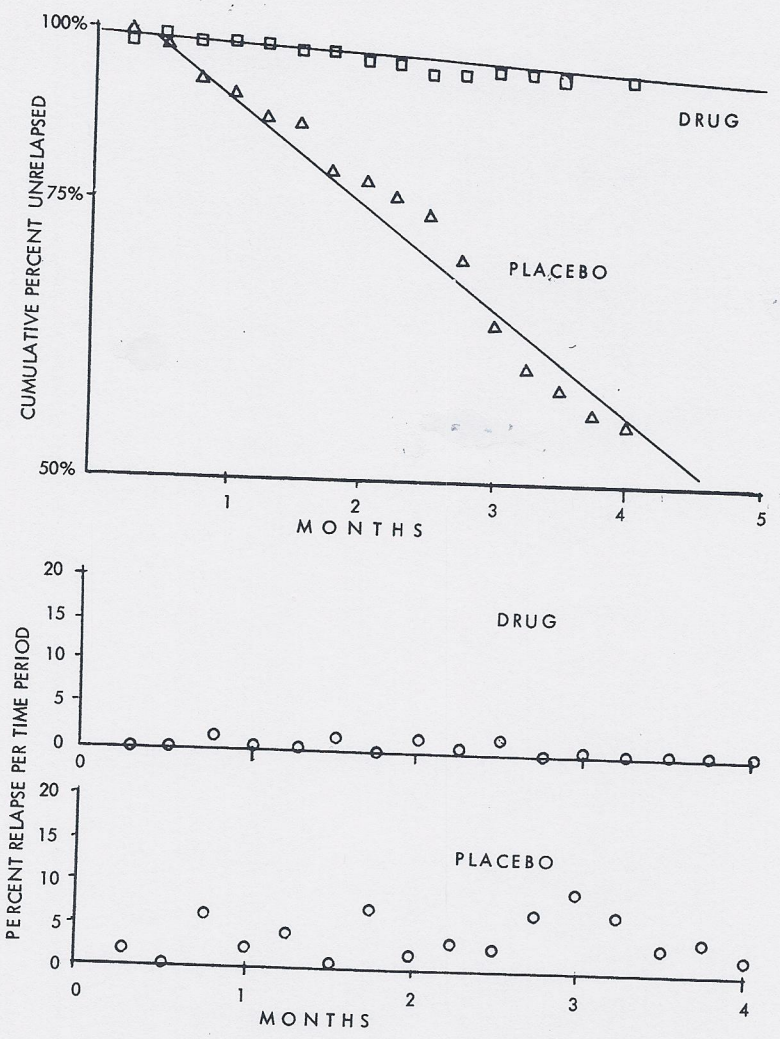


FIG. 3. The survival distribution is expressed as the Ln_e of % of patients unrelapsed over time. Data: Caffey et al. (1). Bottom panel: Hazard function, same data.

and begin to catch up to the placebo patients. Thus the drug-placebo difference narrows. Consequently, the drug-placebo difference observed would depend in part on the length of the follow-up period. It should now be apparent that the absolute percentage of patients relapsing at a given arbitrary time point is not a valid measure. To make the point in a slightly different manner, we calculated phi coefficients for the identical drug and placebo relapse rates with a sample size of 1,000/group and presented a much smoother curve in Fig. 1C.

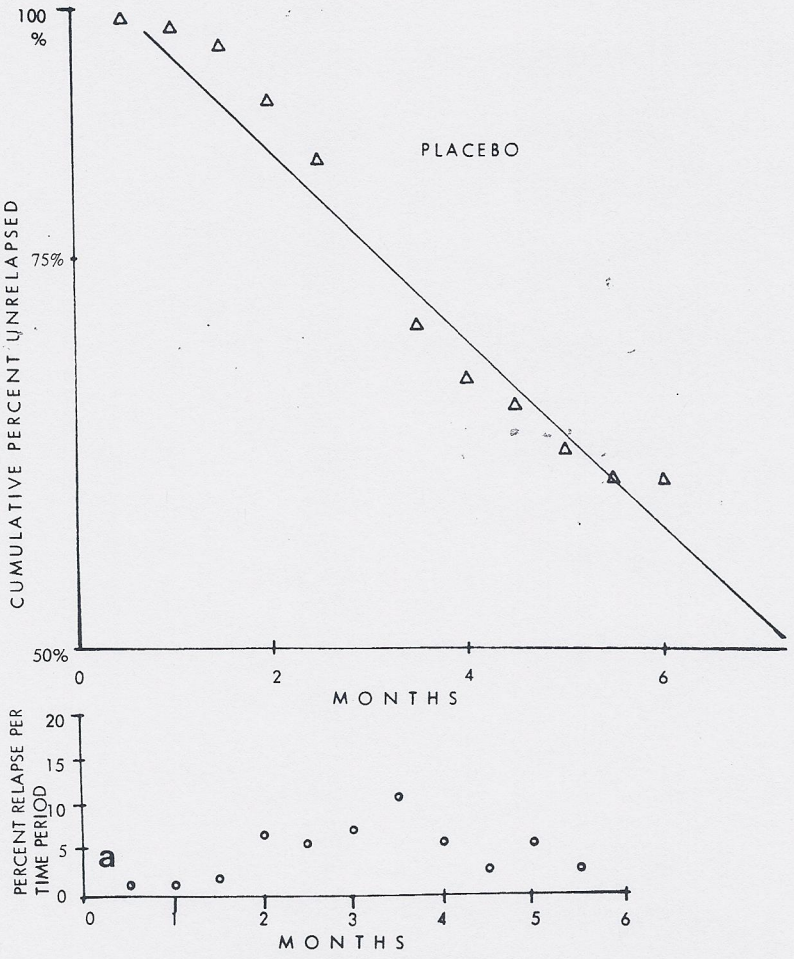


FIG. 4. The survival distribution is expressed as the Ln_e of patients unrelapsed. Bottom panel: Hazard function. Data: Prien and Cole (7).

An alternative statistic that has been used for such an analysis is the mean survival time. This, however, can be calculated only if patients can be followed long enough to know that all at risk for relapse have done so. If the follow-up time is brief, the mean rate could be quite misleading, as many unrelapsed patients may relapse differentially at later points in time.

Let us first consider the case in which one wishes to determine that a drug is more effective than a placebo. We have already pointed out that the chi square statistic which is the commonly used method of analysis is invalid for this purpose. It is excessively conservative in that it is usually too small to serve as an accurate representative presentation of the difference between two

treatments. Indeed, more appropriate ways of thinking about such data were published in 1963 by Halley (4). There are, at present, a number of methods available for dealing effectively with such survival curves. These have been developed in a wide variety of areas concerned with survivor distribution such as insurance policies, reliability of manufacturing goods over time, radioactive decay, or other measures based on concepts of half-life. Both parametric and nonparametric methods have been employed. The latter have the advantage of not assuming a given distribution; we would note that normal distributions are not typically observed in survivor-type data. Examples of two applications of nonparametric methods to such data are those of Kaplan and Meier (6) in 1958 and Gehan (3) in 1965. We have used the method of Gehan to analyze the hypothetical data. His nonparametric method is far more sensitive than a chi square, and of equal importance from a theoretical point of view; it is conceptually a much more satisfactory method for analyzing the data. In sum, the suitable methods for assessing significant differences between two treatments would be those methods employed by Gehan, Kaplan and Meier, and a variety of other survival curve methods. The Gehan method shows a clearly significant beneficial effect of phenothiazine over placebo in preventing relapse of schizophrenia (8) in the data of Wing and Leff ($p < 0.005$).

There is another important practical consideration for why such methodology is superior to the conventional chi square—it can take into account those patients who drop out of the study and do not participate in the full follow-up period. For example, let us say that a 4-year grant is awarded for the completion of a 3-year follow-up. At the start, on an average one patient is added to the study each month. By the end of the year, as the study is progressing, a fairly rapid patient inflow is achieved with about one patient entering the study each week. However, if it is a 3-year follow-up, no patient can be entered after the first year because there would be insufficient time to follow-up the patients by using the conventional chi square methodology. However, by the use of survival distribution methods, patients can be entered into the study throughout the second, third, and even part of the fourth year. This system allows the patients who entered the study for shorter periods of time to be counted in the study. That is, if a patient enters midway during the fourth year and is studied only 6 months prior to the end of the study, he can be included regardless of whether he relapses or does not relapse. In this hypothetical example, let us say that 6 patients were entered in the study during the first 6 months, 10 the next 6 months, and then 24 patients a year entering regularly for the next 3 years. Let us also say that those entered in the third year would have been followed for anywhere from 1 to 2 years, whereas those entered during the first year could have been followed anywhere from the third year to the fourth year. Had the chi square methodology been used, only 16 patients could have been studied. However, by using the more powerful method described, it would become possible for 100 patients to be studied. In these days of time-limited research grants, such an advantage is of considerable importance. Finally, an answer

as to whether a drug is effective can be obtained in a much shorter length of time, thus bringing this clinically important information to the general practitioner more rapidly.

A further reason for plotting the survival curve is that it might help in generalizing to other situations or time periods. Projection of a curve can allow a reasonable guess. In fact, making the assumption that 50% of patients who do not relapse at 6 months will never relapse is a projection of a survival curve. If the survival curve is plotted, it is much easier to appreciate the assumption made in projecting this curve for a longer period of time. Examining the survival distribution allows one to think more clearly about extending distribution curves hypothetically over a longer period of time. A plot of the curve is important for two reasons: the first is that it yields a more intelligent depiction of the data because it can be looked at more easily; the second, and equally important, reason is that it makes explicit one's assumption underlying such projection of data.

The most important part of this chapter is that the chi square 2×2 table method of thinking allows the unintended projection of an implicit survival curve in the statement that 50% do not relapse with placebo and, therefore, will never relapse. We feel that this is not a reasonable projection and that survival curves explicate the assumptions involved in this or other projections. The long-term study of Hogarty and Goldberg verifies that patients continue to relapse after the 50% point.

We feel that the next appropriate step after having plotted the data for description as survival functions and significance testing is to look at the curve in a quantitative sense, with an eye toward guessing as to mechanism. A logical point of departure is to inquire into the consistency of the rate of relapse per unit time. The answer is that the rate of relapse in most of these curves does appear to be constant per unit of time. Indeed, if this is fitted to Ln_e percent of relapse per unit time, as we did by the least squares method, the fit is really quite good. For example, in the Prien placebo data the correlation coefficient is 0.98 ($r^2 = 0.96$) (relapse rate 10%/month). Similarly, for Caffey's data, the correlation coefficient is 0.98 ($r^2 = 0.97$) for the group maintained on placebo (relapse rate 14%/month). In Hogarty's drug data, the correlation coefficient is 0.96 ($r^2 = 0.93$) for placebo group (relapse rate 7.1%/month), and for the drug group the correlation coefficient is 0.99 ($r^2 = 0.97$; relapse rate 2.6%/month). From a practical point of view, a clinician may wish to know if a patient has received a placebo and, if he has not relapsed for 6 months, whether he can be considered recovered; alternatively, should the drug be discontinued, or should it be administered for a longer period of time. If one projects the relapse rate for the first 6 months to the next 18 months, then one would predict more relapses. This is essentially what was done in the study of Hogarty which verified that the rate observed by Prien and Caffey in the first month does continue for the next 18 months. The clinician will naturally wonder about what will happen in a longer follow-up period. In addition to helping one make

an estimate, such a curve makes one more aware that one is making a projection. Upon inspection of the placebo group in the Hogarty and Goldberg study, it seems reasonable to note that there is curvature apparent in the line at 16 to 22 months (the end of the study), suggesting that the relapse rate may be slightly decreasing with time. It may be that the constant relapse rate applies only to the initial phases. Since the exponential function fits very well (r^2 in high nineties), it provides a reasonable approximation of describing the relapse survival function for the first 18 months. It may not apply to time periods greater than 18 months. It may be that there are two types of schizophrenic patients, one of whom relapse at the rate of approximately 15% a month and a small population of good prognosis schizophrenics who never relapse. Can one resolve this relapse-rate curve into two components: one component of a relatively rapid relapse rate of approximately 15% per month and a second component which either has a very slight relapse rate or a relapse rate of zero? The rate constant seems to be significantly decreasing. We will deal with mixed model curves in a later paper. We will also discuss the use of covariates in survival curves in a longer treatment. This is obviously important in adjusting for a prediction of patients who never relapse.

One of the purposes of empirical data is to try to elucidate the mechanism underlying a phenomenon. For example, let us say that one wants to determine how long an automobile will continue to function. Cars are run on a test track and one discovers that eventually all cars fail either at 400 miles or at 800 miles. This provides a clue to the breakdown because all engines fail at approximately the same time, after having run for either 400 or 800 miles. This provides another clue because the cars had 20- or 40-gallon tanks and operated 20 miles per gallon. On a more generic level, the clue suggests that an individual component failed in certain cars always at 400 miles and in other cars at 800 miles. This focuses the attention on what failed at just these points in time (or miles). Similarly, the shape of the survival curve of patients who relapse following placebo may provide clues. Patients seem to relapse at a constant rate for the first 12 to 18 months. This phenomenon needs explanation. Whatever causes relapse must be such that it produces a constant relapse rate.

Another purpose of the survival curve is for descriptive purposes to give the clinician a feel for the data, a "bottom line" summary. For example, the knowledge that about 10% of patients relapse per month is a simple, but accurate summary of the first 18-month relapse data.

Much of our argument is just sketched in due to space limitation. Data analysis is imprecise since it is based only on graphs presented by the authors (1,5,7), and exact data are needed for more precise analysis. Also, statistical treatment is not necessarily optimal. For example, r^2 is not an optimal statistic for goodness of fit. We plan a more complete exposition of the concept in a future paper. What we emphasize is that the chi square 2×2 table method used in almost all the previous studies is not optimal for the analysis of relapse data, and we suggest the more appropriate methodology, namely, survival curves.

ACKNOWLEDGMENT

This work was supported by NIMH Grant 30938–02.

REFERENCES

1. Caffey, E. M., Diamond, L. S., Frank, T. V., Grasberger, J. C., Herman, L., Klett, C. J., and Rothstein, D. (1964): *J. Chronic Dis.*, 17:347.
2. Davis, J. M. (1975): *Am. J. Psychiatry,* 132:1237–1245.
3. Gehan, E. A. (1965): *Biometrika,* 52:203–223.
4. Halley, E. (1693): *Philos. Trans. R. Soc. Lond.* [*Biol.*], 17:596–610.
5. Hogarty, G. E., and Goldberg, S. E. (1973): *Arch. Gen. Psychiatry,* 28:54.
6. Kaplan, E. L., and Meier, P. (1958): *J. Am. Statis. Assoc.,* 53:457–481.
7. Prien, R. F., and Cole, J. O. (1968): *Arch. Gen. Psychiatry,* 18:482–495.
8. Wing, J., and Leff, J. (1971): *Br. Med. J.,* 2:599–604.