

Distribution of episodes of mental illness in general practice: results from the Second National Morbidity Survey

N C SMEETON

From the General Practice Research Unit, De Crespigny Park, Denmark Hill, London SE5 8AF

SUMMARY The Second National Morbidity Survey, conducted in England and Wales between 1970 and 1976, contains a unique body of information on episodes of mental illness experienced by individuals registered in a representative sample of general practices around the country. This information is used to construct the episode distribution among the individuals surveyed. The Poisson and negative binomial distributions are then used to model the episodes. The Poisson model gives a very poor fit but the negative binomial model is found to fit the data very well. Deviations of the observed data from this model are discussed. The possibility of applying this model at the local practice level is then considered.

It is now well established that general practitioners in Great Britain come into contact with the majority of patients with psychiatric disorders.¹ It is therefore of interest to have an estimate of the distribution of episodes of mental illness as found in the community from the records of general practitioners. Not only would this make the medical profession more aware of the needs of the population as a whole, but the various mathematical models could then be fitted to test hypotheses as to how such a distribution could have arisen. An inspection of discrepancies between the observed and predicted values for various numbers of episodes in a fixed period of time could lead to further refinements of the model.

The first study of this kind² looked at the frequency of minor-injury accidents sustained by women workers in munition factories. Greenwood and Yule³ showed that if all individuals were equally liable to have accidents randomly in time, then the numbers of accidents experienced by a group of workers in a fixed time period should follow the Poisson distribution. (For details of this distribution see Johnson and Kotz.⁴) However, where the individual liability to have accidents varies, the negative binomial distribution is a more appropriate model (see Johnson and Kotz,⁴ and the Appendix). This work was first applied to episodes of sickness in a longitudinal study of shunters working for South African Railways.⁵

A recent nationwide study which collected data from individuals on episodes of illness in general

practice was the Second National Morbidity Survey.⁶ Kilpatrick⁷ calculated the distribution of episodes of all illnesses recorded in the first year of the survey and attempted to fit the negative binomial distribution. However, no work has been done specifically on the distribution of episodes of mental illness among individuals in the community. As this information is available from the Second National Morbidity Survey, it is possible to construct such an episode distribution for those individuals studied.

The Second National Morbidity Survey

This survey began in November 1970 and involved the Royal College of General Practitioners, the Office of Population Censuses and Surveys (OPCS), and the Department of Health and Social Security. A representative sample of practices throughout England and Wales was involved. The survey asked for the routine recording of patients' morbidity as and when they made appointments. An attempt was made to define different types of episode of illness, but it was up to the general practitioner to decide whether or not a new episode had occurred.⁷ In the subsequent analysis, all types of episodes were combined.

Some of the participating practices dropped out of the survey after one or two years. From the practices that continued to record consultations throughout the whole of the survey period of six years, a complete record for six years was obtained on about

Distribution of episodes of mental illness in general practice: results from the second National Morbidity Survey 131

60 000 individuals. Many more joined or left during the survey period, and these practices obtained about 43 000 psychiatric records altogether. For each psychiatric record, the number of episodes has been found for each of the six years, and this information has been stored in a longitudinal file by the OPCS. Such data, therefore, present a unique opportunity for estimating the episode distribution of mental illness over the country as a whole from measurements taken over a relatively long period of time.

Method

Patterns of episodes of psychiatric disorders in patients from the 22 general practices that participated throughout the whole survey were obtained from the longitudinal file containing these data.⁸ It was considered appropriate to study only the disorders in which the usual pattern is of episodes of illness interspersed with periods of remission. Therefore only records of affective psychoses, anxiety neuroses, phobic neurosis, depressive neurosis, physical disorders of presumably psychogenic origin, neurasthenia, insomnia, and tension headache were taken. Chronic disorders such as mental retardation were not included.

A most important consideration is that of equality of exposure. The inclusion of individuals who drop out of a fixed period study as if they had participated fully will bias results in two ways: their time in the study will be less than for those who completed the period, so fewer episodes might be expected, and those who drop out may have different characteristics from the stayers. Therefore only individuals who were in the study for the complete six years were taken. Since for most individuals episodes of mental illness are relatively infrequent, it was felt appropriate to study the total number of episodes over the six year period. Therefore, for each individual, the total number of episodes of all the relevant disorders over the six year period was calculated.

The Poisson and negative binomial distributions were fitted to the derived episode distribution in order to determine whether individual episode liability was approximately constant or varied.

Results

The distribution of the total numbers of episodes of mental illness for the 60 000 individuals studied for the six years is shown in the table. Note the large proportion (75.7%) having no episodes. Only a small proportion (0.9%) average one or more episodes per year. Nevertheless the distribution is very skewed,

Observed distribution and fitted Poisson (P) and negative binomial (NB) distributions for the number of episodes of psychiatric illness in all individuals in the Second National Morbidity Survey for the whole six years

No. of episodes	Observed	(P)	(NB)
0	45067	36944.7	45060.8
1	7917	17634.1	7959.2
2	3256	4208.5	3209.8
3	1550	669.6	1537.0
4	762	79.9	794.0
5	434	7.6	428.2
6	243		237.4
7	127		134.2
8	77		76.9
9	50		44.6
10	26	1.6	26.1
11	15		15.3
12	10		9.1
13	5		5.4
14+	7		8.0

some individuals having more than 14 episodes in the six year period.

From the table it can be seen that the Poisson distribution fits the distribution of episodes very badly. There are far more cases having a large number of episodes than predicted by the Poisson distribution and there are also a larger number of individuals free from episodes. However, the negative binomial distribution seems on the surface to provide a very good fit. The parameters of the negative binomial distribution are $m=0.4773$ and $k=0.2804$, as estimated by the method of maximum likelihood⁹ (see Appendix). The χ^2 goodness-of-fit test gives $\chi^2=4.878$ whereas $\chi^2_{12,0.95}=21.0$, so in fact the fit is excellent.

There were more individuals with no episodes than predicted by the negative binomial model but a shortfall of those with one episode and an excess of those with two episodes (see table).

Discussion

From the analysis it can be seen that the Poisson distribution provides such a poor fit to the episode distribution that some basic assumption used in the modelling is entirely wrong. In fact, judging from the larger than expected numbers of individuals with either none or more than three episodes, the assumption that all individuals are equally liable to episodes in a given period of time seems questionable. It appears that many individuals are very resistant to mental illness whereas others seem to be very prone to illness episodes. Some varying distribution of liability to have episodes among the individuals is implied.

The negative binomial distribution is based on the assumption that individuals experience episodes randomly in time but with varying degrees of liability to have episodes, the underlying liability in the

population following a gamma distribution.³ Although this appears to be a rather restrictive assumption, in fact the gamma distribution is very flexible and can take a variety of shapes if the parameter values are adjusted.¹⁰

In fact the analysis shows that the negative binomial distribution is a very useful model for describing the episode distribution of mental illness over England and Wales as a whole. The deficit in the number with one episode and the excess in those with none or two is in line with the observations of Kilpatrick,⁷ although, presumably due to larger numbers in his sample (about 315 000), Kilpatrick found the negative binomial model rejected by the χ^2 goodness-of-fit test. Kilpatrick made two suggestions regarding these deviations. Out of date registers will include individuals no longer alive or in the area, who will thus have no episodes. Additionally, the recording of multiple diagnoses by the general practitioners at one consultation may yield increased numbers with two or more episodes. This will be particularly so if a separate episode is recorded for each diagnosis at a single consultation, as each recurrence of illness could then lead to several episodes being recorded rather than just one. It may be, however, that the excess in those with zero observations could be due to a reluctance on the part of the individual to approach a doctor about a problem which might carry a social stigma.

That individuals across England and Wales are not equally liable to be recorded as having episodes of mental illness is clear from the results. The episode distribution gives an idea of how mental illness is likely to be spread across the country. In terms of application, however, it is of interest to see whether these findings can be used by the general practitioners in order to predict future local mental health needs. Unfortunately, these results could be very misleading if applied directly because much of the variation in episode experience could be due just to the effects of age and sex as well as differences between practices in recording numbers of episodes. Dunn⁸ has shown that these three factors are all important in describing the variation in anxiety and depression observed in this survey. The between practice variation in rates of diagnosis of disorders is particularly large. Dunn¹¹ remarks that in one practice almost 10% of the registered patients were diagnosed as suffering from affective psychoses at least once during the six years of the survey whereas in most practices there were almost no records at all. For the more common disorders the variation is just as disturbing—from 3.0% to 22.2% for anxiety and from 2.7% to 21.9% for depression, for instance.

Due to large inter-practice variations it is impossible to tell from the overall episode

distribution whether the negative binomial model is applicable at the level of the local general practice. Even if this were the case it is almost certain, from the inter-practice variation already noted, that the individual practice values of m and k would be very different from the aggregate values. In other words, both the mean (location) and the tailing off (spread) of the distributions may vary widely from practice to practice even if the same basic model is applicable.

The episode distributions therefore need to be found for the individual practices in the study. If the negative binomial model is applicable, the effects of age and sex can then be looked at to see whether, after accounting for these, any of the variation could be due to variation in individual proneness. Only at this stage will the local general practitioners be able to begin to search for high risk groups in order to predict the future needs of his patients.

In conclusion, the negative binomial distribution does seem a reasonable model for the variation over the country as a whole, but further work will be required in order to assess its usefulness at the level of the local general practice.

The author is supported by a grant from the Department of Health and Social Security to the General Practice Research Unit.

Correspondence and reprint requests to: Nigel C Smeeton, Statistician, General Practice Research Unit, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF.

References

- Shepherd M, Cooper B, Brown AC, Kalton GW. *Psychiatric illness in general practice*. London: Oxford University Press, 1966.
- Greenwood M, Woods HM, Yule GU. A Report on the Incidence of Industrial Accidents upon Individuals with Special Reference to Multiple Accidents. Report 4, Industrial Fatigue Research Board, London, 1919.
- Greenwood M, Yule GU. An enquiry into the nature of frequency distributions representative of multiple happenings, with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *J R Statist Soc* 1920; **83**: 255–79.
- Johnson NL, Kotz S. *Distributions in statistics*. Volume 1, Discrete distributions. New York: Houghton Mifflin, 1969.
- Adelstein AM. Accident proneness: a criticism of the concept based upon an analysis of shunters' accidents. *J R Statist Soc Series A* 1952; **115**: 354–410.
- Royal College of General Practitioners. Second National Morbidity Survey. *J R Coll Gen Pract* 1980; **30**: 547–50.
- Kilpatrick SJ. An empirical study of the distribution of episodes of illness recorded in the 1970–71 National Morbidity Survey. *Applied Statistics* 1977; **26**: 26–33.

Distribution of episodes of mental illness in general practice: results from the second National Morbidity Survey 133

⁸ Dunn G. Longitudinal records of anxiety and depression in general practice: the Second National Morbidity Survey. *Psychol Med* 1983; **13**: 897-906.
⁹ Anscombe FJ. Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika* 1950; **37**: 358-82.
¹⁰ Mood AM, Graybill FA. *Introduction to the theory of statistics*. 2nd ed. New York: McGraw-Hill, 1963.
¹¹ Dunn G. Records of psychiatric morbidity in general practice: The National Morbidity Surveys. *Psychol Med* 1985; **15**: 223-6.

Appendix

Let r represent the number of episodes of illness reported to a general practitioner by a person

registered with that doctor in a given period. If r follows the negative binomial distribution then

$$p_r = \frac{(k+r-1)! k^k m^r}{(k-1)! r! (k+m)^{k+r}}$$

where $r=0,1,2, \dots$ is the number of episodes, p_r is the proportion of persons reporting r episodes, m is the mean number of episodes per person, and k is a further parameter which along with m determines how quickly the distribution tails off. For large values of r the tailing off factor is approximately given by $\frac{m}{m+k}$. In our case, for r greater than four,

$p_{r+1} \approx 0.6 p_r$
 indicating a long, skewed distribution.