SPECIAL ARTICLE

Statistical methods applied in microbiology and epidemiology

Those who work in any area of public health are exposed to large amounts of data and processed information. Some acquaintance with statistical methods can make that exposure more profitable. The objectives of this paper are to describe a suitable framework for governing scientific investigations which involve applied statistics, and to set down guidelines for research workers preparing material for publication. Final sections suggest further reading and give a brief description of some of the statistical software available to assist analysis.

STATISTICS AND STATISTICAL METHODS

Statistics in the form of gathered. numerical data are widely published and provide much relevant information for medical and social studies and health planning. The evolution of the gathering and publishing of these has been described by Galbraith [1]. Current publications of tables and information are listed in the Guide to Official Statistics [2].

The science known as 'statistics' deals not only with collecting data but with their analysis and with making inferences from these analyses. Statistical methods of data processing have grown from standard numerical procedures and statistical theory has developed as a specialized branch of mathematics. Effective practical application combines common sense with an objective approach to the interpretation of data.

PLANNING INVESTIGATIONS

Research workers often say that they need help with their statistical work. A study done in America. initially aimed at discovering computing requirements, found that statistical help was a higher priority and that the greatest need was for help with the planning of their research [3] so as to yield sound definitions and methods of data collection. on which valid (statistical) inference could be based.

The foremost requirement of any study is the formulation of objectives. These should include the specific objectives of the immediate project and the overall objectives to which the present project is a stepping stone. For example. an initial study of an outbreak of gastroenteritis may aim to find a common factor and postulate the source of an outbreak, which might turn out to be a particular restaurant. By the time this is achieved there may be further microbiological evidence to help direct investigations. A second epidemiological study would then be needed to detect which meals and food items were involved. Throughout, the

18-2

common objective would be to find the source and terminate the outbreak and prevent recurrence.

Second, consideration must be given to the type of study to be undertaken. Broadly there are three possibilities: analysis of routine data, observational studies and experiments.

Third, what data are to be collected and from whom; how are they to be gathered, recorded and checked; what is to be done about non-responders, missing data, inconsistencies and errors? In what form must the data be recorded for analysis?

Fourth, what is the required size and cost of the study and are there any ethical constraints?

Pilot studies

Almost all investigations benefit from a pilot study to test thoroughly all aspects of the study design and feasibility. This may be as big as 25% of the final study, as suggested by Deming [4], and if successful these results can be incorporated in the main study. This pilot phase must not be confused with preliminary enquiries used to generate the hypothesis for a study. For example, if a cluster of unusual cases requires an outbreak investigation then a few patients may be interviewed in depth. If a possible common source of exposure emerges from these interviews, then an objective study should be set up to investigate whether this was the actual source of infection. However data from those early patients should *not* be included as it could be argued that they bias the results. Data from a true pilot study can be used to plan summaries (tabular and graphical), and analyses of the main study. This may speed up the final preparation of results. A study using routinely collected data may need no pilot study.

Ideally much time and effort will be devoted to these planning stages, and this should certainly be the case for studies of microbiological methodology and epidemiological research other than urgent outbreak investigations. The latter need speed in order to 'remove the handle from the pump' (see the work of John Snow [5] and a finely judged balance must be struck between speed and reliability.

DATA AND ANALYSIS

Routinely collected data

Analysis of routine data has a long history going back to John Graunt who used the London Bills of Mortality to show, in the seventeenth century, that mortality was higher in infants than adults, and in rural than urban areas [6]. A recent example is a study of deaths in England and Wales which showed a relative increase in death rates among young, unmarried men which was attributed by the author to the human immunodeficiency virus [7]. The data were from routine death certifications and so the causes of death were those transcribed from the death certificate to the death entry by the local Registrar. The observation of excess mortality raised the hypothesis that AIDS or other manifestations of HIV may be the cause, even though the numbers greatly exceeded the AIDS deaths known to the voluntary surveillance scheme. AIDS or HIV were frequently not

mentioned on the 'excess' death entries, either because the HIV infection was not known or it was not thought to be relevant, or because it was wished to avoid distress to the families.

To undertake studies of routine data it is essential to have a thorough knowledge of the data, how they are collected and the errors or biases that can affect them. The results may not be conclusive in themselves and a further study may be required to explore the hypotheses thrown up by the initial study. This is no reason to discourage researchers from publishing work which only raises questions. Such papers can spur others to ask similar questions in other data sets and to seek explanations.

Observational studies and experiments

The essential contrast between an observational study and an experiment lies in the way 'experimental units' are allocated to the 'treatment groups'. In an observational study we use the naturally occurring allocation process and observe the outcome. Thus the health and mortality of men with a history of coal mining may be compared with those of men living in the same area who never go into mines. If adverse observations among the mining group were to be ascribed to occupation then it would have to be assumed that the young men choosing a career in the mines were as strong and healthy as the others. A prospective study of school leavers showed this not to be correct and that, on average, physiologically weaker boys went into mining. Thus in an observational study there is always the possibility that observed differences may be due to some unknown or unsuspected 'confounding' factor.

In contrast, in an experiment the different treatment units (e.g. blood samples, patients, herds of cattle) are randomly allocated to the different treatments (e.g. serological methods applied to blood samples, vaccines given to patients, antibiotics fed to herds). Randomization should ensure that any observed differences between groups are due to the treatments. This does not mean that no effort should be made to ensure that treatment groups are as similar as possible in the light of common sense. Thus, in a clinical trial, patients with similar severity of disease or of similar age group should be equally distributed to the treatment groups. However, at the final stage of allocation randomness must be applied using a mechanical random process, such as tossing a coin or using a computer generated random number, or random identification which cannot be influenced by a human agent. 'Randomly allocated' should not be used as a euphemism for arbitrarily allocated.

For a more detailed description of the differences between, and design of, experimental and observational studies see the work of Galbraith and Palmer [8].

Despite the inherent drawbacks of observational studies, they can be important for preliminary investigations and essential for problems in which experiments are impossible. The latter will include situations such as outbreak investigations and studies of patterns of infection among farm animals. The potential problems of interpreting differences between groups (the 'confounding' factors) can sometimes be anticipated or prevented by the study design and the data collected. Observational studies divide broadly into two groups – retrospective and prospective.

In a retrospective study one starts with the outcome (e.g. infected or not infected) and compares the histories of the two groups to identify causal factors. If one outcome is more readily identified (e.g. the infected) then a case-control study design can be designed by finding uninfected control subjects which are matched for possible confounding factors. If a whole population is being studied (e.g. a retrospective cohort study of a school population) then the histories should include data on potential confounding factors so that they can be accounted for in the analysis. Retrospective studies have the advantage of being fairly simple, easy, inexpensive and fast. But problems may arise if retrospective information is unobtainable and responses to questionnaires may be biased by knowledge of subsequent events (or mistaken impressions from the media!)

Prospective cohort studies (but not retrospective cohort studies) observe a group of subjects and follow them up to see the outcome. They avoid the disadvantages of collecting data retrospectively but subjects may be lost to follow-up, the studies are often lengthy, costly, and may be overtaken by events. They are more often used in chronic than in communicable disease epidemiology.

Data

The quality of any study depends on the data. In most investigations interest centres not so much on the observations themselves as on the group they represent. Thus the observations are viewed as a sample (ideally random) from some population. For example, in a study comparing test methods for detecting HIV antibody the samples should represent the infected population. From this point of view non-representation of large parts of the population (or, in another example, non-response in a survey) can result in serious lack of information. In surveys of the prevalence of tuberculosis it was well known that the disease was much more prevalent among people who were unwilling to be X-rayed. Recent surveys of HIV have shown that seroprevalence is often higher among patients who do not consent to be tested than among those who do [9]. Such potentials for bias must be acknowledged in any conclusions.

The use of microcomputers to enter and check data as they become available can greatly enhance the quality of a study. Missing data, errors and inconsistencies can be identified rapidly so that there is the best possible chance of making corrections, which might not be the case if data processing does not commence until all results are assembled.

Tabulations and graphs

The first analytical step is to prepare simple tabulations of each variable, or, in an outbreak investigation, comparing cases and non-cases variable by variable. Such tabulations have two great virtues. They may reveal discrepancies in the data which must be corrected and they usually reveal the essential messages which the data contain. While formal inference proceeds from hypotheses via observation and experiment to new hypotheses, it is important to stand aside, as it were, to look and listen to what the data have to say. Graphical displays, which can readily be constructed with modern software, can assist in this. This first stage should never be omitted.

Analysis

In some situations the simplest tabulations may be all that is necessary, but frequently more analysis is required to bring out or confirm the essential features of the data. Analysis is both theoretically and practically analogous to filtering radio waves to obtain clear messages concealed by a jumble of irrelevant noise in the raw signal or data. Calculating an average corresponds to passing the data through a simple filter after which the population mean emerges plus some minor residual error – the scale of which is measured by the appropriate standard error. More sophisticated analysis may enhance the signal.

Informed assessment can be made only if the data have been analysed critically, often from the position of 'devil's advocate'. For example, analysis of food consumption histories in a community where there has been an outbreak of food poisoning can throw up spurious associations. If the illness is predominantly among the young then all foods favoured by this age group can appear to be candidates for the guilty vehicle of infection. It is therefore essential that analysis first takes into account age. Critical assessment of correlations must always seek to separate direct from indirect associations, and acknowledge doubt if this is not possible. Any disease with a seasonal pattern in its epidemic curve will show significant correlation with meteorological variables, but this is not to say that it is cause and effect, with weather influencing spread. This will be suspected only if the correlation remains after the seasonal pattern has been taken into account, as is illustrated with total mortality in England and Wales which has been closely studied to estimate the effects of influenza epidemics and weather. Deaths have a strong seasonal pattern, with peaks in the winter period, but if the temperatures fall lower than the average for that time of year then deaths have been shown to increase correspondingly, and thereby demonstrating the direct influence of temperature on mortality [10].

It may be easy to measure a large number of variables in observational studies and experiments. One is therefore tempted to embark on multivariate analysis to explore how all the various factors interact. Twenty years ago the computations involved were usually prohibitive, but now they can be done on a desk top microcomputer. However, there are serious problems which must be faced when attempting multivariate analysis, and which are seldom discussed in users' guides to statistical computer packages.

The first is the size of multidimensional space. For example, suppose you have a cohort study of 100 people of whom 30, say, had the infection and there appeared to be age and sex effects which needed to be taken into account when looking for potential risk factors. Even if the subjects are divided into two age groups, split at the median, and exactly half are males then there are only about 25 people to study in each age/sex group. In effect each 25 then have to be sub-divided according to exposure or non-exposure to the risk factor, and the two attack rates compared. The statistical method will in fact combine the evidence within each age/sex group, but clearly the data are being stretched quite thinly over just three variables even with a sample size of 100.

A multivariate analysis aiming to explore the effects of many variables is essentially a large numbers game and should use as large a data set as possible.

There may inevitably be areas in the multivariate 'space' which are empty of observations and the researcher should be aware of this and what effect it has on the conclusions. In taxonomy this may be a positive help – for example, the observation that certain sub-species never react to a certain test. In other situations it can restrict decisions. For example, in a study of levels of antibody in populations of mixed *vaccinated*, *partially vaccinated*, and *unvaccinated* people, the effects of vaccine batch and time interval from last dose can be thoroughly evaluated only if there is a good spread of observations.

Inference

For the present purpose we wish to separate two areas of statistical techniques. The first is the calculating and displaying of associations found from the data. The second is interpreting, or making inferences, from the calculations. In practice the two areas are intimately entwined.

'Posterior' belief depends on the combination of 'prior' belief plus the data, as Bayes proved over a hundred years ago. The problem is that people's prior beliefs vary and are often ill-formed and consequently different people will draw different conclusions from the same data. Fisher's scheme of 'fiducial' inference and the more widely used scheme of Neyman–Pearson hypothesis testing are designed to circumvent this problem. The researcher can report on whether the 'null hypothesis' that factor A has no effect on outcome B is or is not rejected by the data at a specified level of probability. Such statements are perceived as objective, although both the choice of data (or its method of collection) and the conviction produced by the statistical test result do depend on one's prior beliefs more often than we care to admit. Furthermore, the correct use of Neyman–Pearson theory requires that a specific hypothesis (usually referred to as H_0) is formulated before the data are collected and studied, and is then compared with a specific alternative hypothesis (H_1). Thus the scheme of inference starts with prior beliefs that either H_0 or H_1 is true.

After the analysis one may come to the correct conclusion but, because we are dealing with a probability model, two sorts of error are possible. First we can mistakenly reject H_0 and adopt H_1 when the former is true – this is known as an error 'of the first kind' and the P value of the test indicates the probability of this error and is usually desired to be $\leqslant 0.05$. Second, we can accept H_0 when H_1 is true – this is called an error 'of the second kind'. The tendency for research workers and statisticians not to specify the alternative hypothesis precisely means that second type errors are all too common. Studies are set up which could be certain of detecting only the grossest risks or associations and trials are done with so few subjects that only a miracle would produce a significant result. The danger is that, after a few such studies, a null hypothesis can enter the corpus of 'scientific knowledge' on the basis of no positive evidence whatever. For example, a laboratory method may be changed because it seems as good as the old one when comparability has not really been assessed.

To avoid errors of the second kind the studies must be large enough to ensure that H_0 will almost certainly be rejected if the alternative hypothesis is true, even when the results of the study are, by chance, more favourable to H_0. The probability of correctly rejecting H_0 in favour of H_1 is called the power of the significance test. Power should be considered at the planning stage of a study,

when thought may also be given to practical significance. Very large studies may detect small but statistically significant differences which are of little or no practical importance.

The formal setting up of hypotheses can be stultifying because there are situations where neither hypothesis is true. This has led to the increasing use of confidence intervals in the presentation of such results. These show a range of parameter values with which the data are almost certainly compatible. If the confidence interval includes the null hypothesis, then H_0 is not rejected by the data. But the width of the confidence interval immediately shows whether the study has any chance of detecting a realistic alternative to H_0. Thus if H_0 and any reasonable H_1 are both included in the confidence interval, then this study is too small to be useful.

An example might be an experiment to compare the detection rate of a new diagnostic technique with the current method. The null hypothesis would be that d (the difference between the two rates) is zero and it might be desirable to know whether the new test was capable of detecting $\geqslant 5\%$ more cases among the experimental population. If the study produced confidence interval for d of -3% to $+9\%$ then the null hypothesis could not be rejected, but neither could you exclude the possibility that the test method was slightly worse or appreciably better.

The use of multiple significance tests to test the many associations that may emerge when data are analysed is not strictly included within the formal theory so far described. Such tests can serve as an informal guide to associations that should be investigated in subsequent studies. In particular, if 20 independent 5% significance tests are applied then at least one may be significant by chance. Similarly, 20 independently calculated confidence intervals are likely to exclude H_0 at least once by chance. Statistical theory does provide for multiple significance tests in the general context of multivariate analysis. However, tests of general hypotheses are much weaker than tests of specific hypotheses. So if a specific hypothesis can be formulated *a priori* then this should be done at the planning stage.

When decisions are being made about the conclusions of a study, then power, error and multiple testing should all be considered within the assessment of the statistical part of the evidence.

Non-parametric methods

The theory of inference generally requires that the data can be viewed as a sample from a known distribution – e.g. Normal (bell shaped). Bacterial counts seldom are, but may be so on a logarithmic scale. The essential, initial data summaries should show whether data appear Normal on the appropriate scale. Some bacterial counts, such as total coliform organisms in water samples, often turn out to be skewed right on the original scale but skewed left on a logarithmic scale. In such a situation it may be best to employ statistical methods which do not assume anything about the shape of the distribution – the so-called 'non-parametric' or 'distribution-free' methods which include special tests for correlation and the χ^2 tests on contingency tables. An objection to such methods has been that they provide only tests of null hypotheses. More recently methods have been developed for constructing confidence intervals corresponding to non-

parametric tests, but the computer programs for such work are only just becoming available.

Small numbers (particularly in the context of outbreak investigations)

In an outbreak investigation speed is important if there is to be a chance of halting the outbreak or preventing another. Case-control or cohort studies of the ill and the well need to be thoughtfully planned but implemented quickly. They must be of sufficient size to stand a chance of detecting the cause without being so large that results are delayed. Very often the size is governed by the size of the outbreak which, if small, will necessitate interviewing multiple controls per patient to give a worthwhile sample size (although there is little to be gained by having more than four or five controls per case). It is important to remember that, in the investigation of an outbreak, we are usually seeking one cause and one source. The size of the risk may be irrelevant. Chronic disease epidemiology generally involves multiple causative factors and the researcher will be seeking to identify each risk factor and to measure their relative importance. Text-book equations for estimating relative risk and giving confidence intervals for the estimates are often approximate and rely on large sample sizes, such as are studied in chronic disease epidemiology. Researchers should use more accurate methods when there are less than 20 cases. Some appropriate methodology has been developed over recent years [11, 12]. Large sample methods and most small sample methods can be applied using the statistical computer software listed in the final section of this paper.

Worked examples of methods described by Breslow and Day and applied to outbreak investigations have been published by one of the authors [13]. Small numbers and multiple, variable numbers of matched controls are not uncommon in the work of the Communicable Disease Surveillance Centre. Outbreak investigations can be demanding and are sometimes done under extreme pressure and many hypotheses will be tested during the statistical epidemiological studies. The theoretical weakness of such multiple testing is usually counterbalanced by complementary, independent evidence such as microbiology. Nevertheless, there are occasions when a decision has to be made to issue a health warning or withdraw a foodstuff from sale based on statistical evidence alone. In such a situation it is very important to avoid errors of the first kind. Close attention will be paid to the probability of H_0. One example involved a telephone survey over a weekend. Statistical analysis was done on the Sunday night and the implicated product was voluntarily and effectively withdrawn from sale on the Monday [14]. The subsequent isolation of salmonella from the product confirmed that the correct decisions had been made.

Practical issues in outbreak investigations

In an outbreak investigation there will often be microbiological and other scientific evidence to be weighed as well, which will play a part in the final decision making. The drawing of conclusions, therefore, will not usually be confined to the statistical analysis. Indeed there may be occasions when microbiology demonstrates the source of an outbreak when statistical analysis does not. The dread of any epidemiologist who is investigating a food poisoning outbreak is to trace it

back to a single function and find that there was a set menu. Thus almost everyone present will have eaten every item. The tabulated columns of numbers ill and not ill who did not eat such food will be so small that 'no statistical association' will be found between illness and exposure to any particular food. This is a situation where the report of the investigation should make it clear that the power of statistical tools was so small that nothing conclusive could have been expected. It should not be reported to imply that there was 'no risk' associated with eating each item, rather that it was not a suitable area for analytical statistics. The hope is that the microbiologists will find the relevant organism in left-over food.

An example of how statistical analysis, in a more complicated outbreak investigation, tied in with microbiological and engineering results is the outbreak investigation of Legionnaires' disease associated with Stafford District General Hospital in 1985 [15]. The epidemiological surveys revealed two different patterns. The cases of Legionnaires' disease appeared to have been exposed especially to one part (the Out Patient Department) of the hospital. The large numbers of staff who had not had Legionnaires' disease, but who had acquired legionella antibody, had spent much of their work time in one wing of the hospital. This was the wing which included the OPD but extended to the two floors above. Multivariate analysis was used to demonstrate independent associations with each floor. Engineering and microbiology studies identified two possible routes by which an aerosol contaminated with legionella could have entered the hospital. One route would have contaminated the entire wing but the other was concentrated into the OPD. Evidence from individual disciplines would have been inconclusive, but put together it was possible to hypothesize the history of the epidemic, because results could be pieced together like a successful jigsaw puzzle.

PREPARATION FOR PUBLICATION

Statistical analysis is an essential part of many epidemiological and microbiological studies. If the analysis is planned according to the aims of the study and kept as simple as is sensible, then it should be possible to convey the results to a wide audience. The following points should be considered:

(a) If possible special data should be collected in line with the aims of the study and the statistical analysis should be planned ahead. Even if routine data have to be used, the direction in which the analysis will go should be thought through in advance to ensure objective assessment.

(b) A thorough scrutiny of the data must be made. Tables or graphs should be used to present the actual data or, at the very least, informative summaries. These will set the scene and show whether subsequent analyses are justified. They will also indicate whether the quantity of data will be powerful enough for the investigations required. They should give some indication of the quality of the data – whether there are too many 'not knowns' or under representation in some groups. If any of the sample sizes is very small then it is unrealistic to expect much or any evidence from statistical analysis.

(c) Statistical results should be qualified to illustrate their accuracy. Thus confidence intervals should be quoted for all parameters estimated (using a sensible and not excessive number of digits or decimal places). If it is appropriate

to quote a probability then this should be presented numerically and not grouped into 'significant' or 'not significant'.

(d) The statistical analysis is an aid to decision making. Critical assessment of the data and complementary evidence from other disciplines should be presented to show how the authors reached the final conclusions.

FURTHER READING

The application of statistical methods to medical research was pioneered by Bradford Hill at the London School of Hygiene and Tropical Medicine. His classic text book 'Principles of Medical Statistics' is still worth careful reading [16]. More recent publications include Armitage and Berry [17], Bailey [18], Kirkwood [19] and Breslow and Day [11].

STATISTICAL SOFTWARE

All but the simplest statistical analysis is facilitated by using a microcomputer. Extensive software is available for use with PC machines (e.g. IBM or similar) which use the MS DOS operating system. The STATXACT [20] package is very easy to use and provides exact analyses of contingency tables, confidence intervals for odds ratios and combined analyses of stratified contingency tables. MINITAB [21] handles data tabulated in columns as on a work sheet; calculations generate new columns or summary statistics which can be printed or stored for further calculations. Most standard statistical analyses are included, e.g. t test to compare columns of data, a non-parametric equivalent, simple and multiple regression. It is particularly suitable for relatively small data sets involving less than a thousand rows of up to ten column variables.

EPI-INFO [22] developed at the United States Centers for Disease Control, Atlanta, with WHO support, is distributed free and is an attractive proposition and will handle larger studies. It is used widely in outbreak investigations. It includes three components – handling data entry, statistical calculations and report writing. Data entry can be arranged to correspond to a form or questionnaire so that data can be entered directly from the documents. Data may be verified by double entry (as was done with punch cards), and cross-checked for validity. Large data sets using several forms with complex relationships can be handled. The package includes facilities for the statistical analyses most appropriate to epidemiological studies as well as a word processor for designing forms and creating reports which can incorporate the results of analyses.

HARVARD GRAPHICS [23] provides publication quality figures, when printed on a laser printer. These can be incorporated into text produced by a word processing package such as WORD PERFECT [24].

Many larger studies are analysed with the statistical package for the social sciences, SPSS [25], or the more advanced general package SAS [26], both of which are now available in versions for microcomputers. SPSS includes a data entry section. Alternatively a standard data base or a word processor can be used to create files of data which can then be read into packages already mentioned, or other advanced programmes. However facilities for checking and validating data

vary and may be limited with some data bases. Among the advanced statistical packages the most important are GLIM [27] which constructs regression and logistic regression models for qualitative and quantitative data, and EGRET [28], which implements the analytical methods developed by Breslow and Day [11].

H. E. TILLETT
Communicable Disease Surveillance Centre, Public Health Laboratory Service,
61 Colindale Avenue, London NW9 5EQ.

R. G. CARPENTER
Department of Epidemiology and Population Sciences, London School of
Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT.

REFERENCES

1. Galbraith NS. Communicable disease surveillance. In Smith A, ed. Recent advances in community medicine. No. 2. Edinburgh: Churchill Livingstone. 1982: 127–41.
2. Central Statistical Office. Guide to official statistics, periodic publication. London: Her Majesty's Stationery Office.
3. Harris RW. Final report on computing needs of medical research workers. Part 2 of contract V1005P–9571. California: Systems Development Corporation. 1963.
4. Deming WE. Some theory of sampling. London: John Wiley, 1950.
5. Snow J. Snow on cholera. New York: Hafner Publishing Company, 1965.
6. Graunt J. Natural and political observations mentioned in a following index and made upon the Bill of Mortality. 5th ed. London: J Martin. 1676.
7. McCormick A. Trends in mortality statistics in England and Wales with particular reference to AIDS from 1984 to April 1987. B M J 1988; **296**: 1289–92.
8. Galbraith NS, Palmer S. General epidemiology. In: Topley and Wilson. Principles of bacteriology, virology and immunity. 8th ed. London: Edward Arnold.
9. Gill ON, Adler MW, Day NE. Monitoring the prevalence of HIV: foundations for a programme of unlinked anonymous testing in England and Wales. B M J 1989; **299**: 1295–8.
10. Tillett HE, Smith JWG, Gooch CD. Excess deaths attributable to influenza in England and Wales: age at death and certified cause. Int J Epidemiol 1983; **12**: 344–52.
11. Breslow NE, Day NE. Statistical methods in cancer research. Vol. 1. The analysis of case-control studies. Lyon: IARC. 1980.
12. Mehta CR, Patel NR, Gray R. Computing an exact confidence interval for the common odds ratio in several 2×2 contingency tables. J Am Statistical Assoc 1985; **80**: 969–73.
13. Tillett HE. Statistical analysis of case control studies in communicable diseases. Int J Epidemiol 1986; **15**: 126–33.
14. Cowden JM, Lynch D, Bartlett CLR, et al. A national outbreak of *S. typhimurium* DT124 caused by contaminated salami sticks. Epidemiol Infect 1989; **103**: 219–25.
15. O'Mahoney M, Stanwell-Smith R, Tillett HE, et al. The Stafford outbreak of Legionnaires' disease. Epidemiol Infect 1990; **104**: 361–80.
16. Bradford Hill A. A short text book of medical statistics. 11th ed. London: Hodder and Stoughton. 1984.
17. Armitage A, Berry G. Statistical methods in medical research. Oxford: Blackwell Scientific Publications. 1987.
18. Bailey NTJ. Statistical methods in biology. 2nd ed. London: Hodder and Stoughton. 1981.
19. Kirkwood BR. Essentials of medical statistics. Oxford: Blackwell Scientific Publications. 1988.
20. STATXACT. Cytel Software Corporation. Cambridge. MA, USA. 1989.
21. MINITAB 71. Minitab Inc. State College, PA, USA, 1989.
22. Dean AD, Dean JA, Burton JH and Dicker RC. EPI-INFO Version 5: a word processing, data base and statistics program for epidemiology on microcomputers. Centers for Disease Control. Atlanta, Georgia, USA, 1990.

23. HARVARD GRAPHICS. Software Publishing Corporation, Mountain View, CA, USA, 1987.
24. WORD PERFECT 5. Word Perfect Corporation, Orem, Utah, USA, 1989.
25. SPSS/PC. SPSS Inc, Chicago, Illinois, USA, 1988.
26. SAS 6.03. SAS Institute Inc, Cary, NC, USA, 1988.
27. GLIM 3.77. Generalized Linear Interactive Modelling System, National Algoriths Group, Oxford, UK, 1986.
28. EGRET. Statistical and Epidemiology Research Corporation, Seattle, Washington, USA, 1989.