

On statistical methods for analysing the geographical distribution of cancer cases near nuclear installations

JOHN F BITHELL AND RICHARD A STONE

From the Department of Statistics, University of Oxford, and the CSIRO Division of Mathematics & Statistics, Melbourne, Australia.

ABSTRACT There is great public concern, often based on anecdotal reports, about risks from ionising radiation. Recent interest has been directed at an excess of leukaemia cases in the locality of civil nuclear installations at Sellafield and Sizewell, and epidemiologists have a duty to pursue such information vigorously. This paper sets out to show that the epidemiological methods most commonly used can be improved upon. When analysing geographical data it is necessary to consider location. The most obvious quantification of location is ranked distance, though other measures which may be more meaningful in relation to aetiology may be substituted. A test based on distance ranks, the "Poisson maximum test", depends on the maximum of observed relative risk in regions of increasing size, but with significance level adjusted for selection. Applying this test to data from Sellafield and Sizewell shows that the excess of leukaemia incidence observed at Seascale, near Sellafield, is not an artefact due to data selection by region, and that the excess probably results from a genuine, if as yet unidentified cause (there being little evidence of any other locational association once the Seascale cases have been removed). So far as Sizewell is concerned, geographical proximity to the nuclear power station does not seem particularly important.

There is great public concern about the risks—especially of malignant disease—from ionising radiation. Much publicity in the media is given to anecdotal reports of clusters of cases and the force of public opinion is powerful enough to have a major influence on policy making, even in the face of the scientific opinion of the nuclear establishment.

That there are risks of malignant disease resulting from ionising radiation is well established. A number of epidemiological studies—such as that of the atomic bomb survivors of Hiroshima and Nagasaki¹—permit some estimates of the risk per unit dose to human tissue. These estimates, if applied by extrapolation to the much lower doses likely to be experienced by the general public as a result of routine operation of nuclear power plants, imply increments of risk so small as to be very difficult or impossible to detect. Although considerable uncertainty surrounds the interpretation of low dose studies, evidence slowly accumulates that extra cases of cancer and leukaemia will result from low doses applied to very large populations, such as the considerable number of fetuses exposed to X-rays during obstetric investigations.^{2–4}

When a cluster of cases appears in a village like Seascale, near the Sellafield nuclear reprocessing plant in Cumbria, Northern England,⁵ reactions tend to be

polarised. The experts quite reasonably point out that such an excess cannot possibly be indicative of the underlying risk to the general population on even their most pessimistic assumptions; it must therefore arise by chance (enhanced by the effect of selecting evidence) or be due to some cause or artefact unrelated to the radiation source in question. Those affected not unnaturally take a different view, as also do political interests opposed to the use of nuclear power. They mistrust the calculations and the assumptions on which they are based; the very fact that ionising radiation is undetectable by the human senses engenders fear and the suspicion that there are sources and pathways of contamination that are as yet unknown.

Epidemiology lies at the heart of this polarisation and its practitioners have a strong duty to pursue such anecdotal evidence as vigorously and fairly as possible, both to ensure that public concern does not outrun the actual dangers by an unreasonable margin and also to maximise the chance that any previously unknown risk may be detected. It is the thesis of this paper that the epidemiological approaches and methods that are most commonly used can be improved upon. We begin by examining some general issues.

Epidemiological methods

In clarifying our ideas it will be helpful to consider a number of general distinctions.

TWO PURPOSES OF EPIDEMIOLOGY

Stone^{6,7} draws a distinction between studies in which a risk factor, for example radiation dose, can be quantified beforehand and one in which it cannot. Neither kind of study is like a controlled experiment, but at least in the first kind it will usually be possible to formulate carefully an hypothesis to be tested, while any systematic relationship between exposure and risk will provide valuable corroboration of the causal reality of the relationship investigated. The second kind is far more speculative and has all the drawbacks of "data dredging." At the heart of any inferential process lies the problem of selection: significant results are observed, but it is not at all clear from what set of potential (and probably non-significant) results they have been selected. Nevertheless this is the situation that arises in the investigation of reported clusters. Scientifically one may prefer to accord them no inferential significance; politically and morally this is hardly acceptable.

DIFFERENT KINDS OF CLUSTER

Next we must consider more carefully what we mean by the term "cluster". Some authors prefer to reserve the word exclusively for groups of cases unexpectedly close both in time and space, but this hardly accords with common usage and is arguably not a very useful restriction. A much more important distinction is that between, on the one hand, a group of cases close in time, space or both because of some interactive mechanism such as contagion, and, on the other hand, one which is a concentration of cases due to a locally elevated risk, individual cases occurring independently of one another. The former, interactive, situation has been the subject of much study in cancer epidemiology, partly because of the possibility that an infectious, viral agent might be involved in the induction of certain tumours. There is very little clear evidence of this, but appropriate tests have been intensively studied, especially those that are derived from the work of Knox⁸ on pairs of cases close in time and space. Smith⁹ has reviewed the literature, while Barton *et al.*,¹⁰ and Bradshaw¹¹ have studied the power of this class of tests, their general conclusion being that they can be very powerful against certain hypotheses of contagious spread.

It is not to be expected, however, that Knox type tests would be particularly powerful against alternatives of the locally elevated risk type, which are of more interest in the context of radiation carcinogenesis. Here the underlying mechanism will

generate a non-homogeneous Poisson process in which cases occur at random in space and time, though not at the same rate or density, and the best tests in such a situation must presumably be based on the information individually contributed by the individual cases. Of course the most suitable test will still depend on the kind of alternative envisaged, but it seems unlikely that anything will be gained by considering pairs of points, as in Knox type tests.

SUSPECT FOCI

The final general distinction we wish to make is between the situation in which there is a suspect focus or source of risk and that in which there is not. This distinction is not quite as clear cut as may appear at first sight. An investigation in which a source of risk is postulated in advance of observing any data is obviously inferentially ideal, but not all that usual. At the other end of the spectrum a "cluster" near no previously suspected focus will typically draw attention to a variety of possible explanations. In between lies the all too familiar case: a cluster is reported both because of apparently high disease frequency and also because of the proximity of some unpopular or suspect installation.

The distinction is clear cut as far as possible methods are concerned, however, and we shall in this paper concern ourselves exclusively with the case where a focus of risk can be identified. This focus will typically be a point on a map, though other possibilities exist. At all events we must assume that this focus is specified a priori while recognising that, to the extent that this is not so, the inference must be weakened to a degree which is very hard to quantify.

Conventional methods of analysis

The traditional approach to analysing data suggesting an elevated risk near a focus F is to examine incidence rates for sub-regions around F, comparing them with national rates and determining in an ad hoc fashion whether there is any obvious locational pattern in the rates. Thus, after the Yorkshire Television programme on Sellafield, for example, Craft and Birch¹² and Gardner and Winter¹³ examined registration and death rates for malignant disease in the surrounding area.

While this approach has the merit of being the "obvious" thing to do and of being easily comprehended, it has a number of drawbacks:

- (i) Although the selection of the regions will be determined largely by the availability of population data, there will nevertheless be some arbitrariness in how much to aggregate them and how large a total region to consider.

- (ii) Reliance on an ad hoc appraisal of the locational relationship between F and elevated rates introduces an element of subjectivity.
- (iii) If rather large sub-regions are chosen there is every likelihood that a small genuine locally elevated risk will be lost by incorporation into a greater whole showing no elevation. This may explain what happened in the registry based analyses of the Sellafield data, which were criticized by the makers of the television documentary.¹⁴
- (iv) On the other hand conventional analyses using sub-regions with very small expectations reject the null hypothesis less frequently than they should by virtue of the discreteness of the data.¹⁵ Thus in a further analysis of the Northern Region Cancer Registry data, Craft *et al*¹⁶ looked at the Poisson counts in the electoral wards in the North of England. Although Seascale had the smallest p value, it is noteworthy that only 19 of the 675 wards were significant at the 5% level, a deficit largely due to the conservatism of exact tests based on calculating the tail probabilities in the Poisson distribution.
- (v) An effective analysis depends on complete ascertainment of cases. Use of data from a population based registry is appropriate if comparisons are made using global rates similarly obtained, but a registry may miss genuine cases appearing in an anecdotal report. If all cases have the same chance of being missed, the registry based analysis will then be unbiased but less than convincing to those personally familiar with the cases omitted.
- (vi) Likewise the calculation of rates and expectations requires knowledge of population sizes which are often troublesome to obtain or even to estimate.

These are familiar problems in epidemiology. None is such as to vitiate the conventional approach completely and it would be unrealistic to hope to avoid all the associated difficulties. Nevertheless these are the problems we seek to mitigate in the rest of this paper.

Locational methods

The problems described in (i) and (ii) above result largely from a failure to find a formal way of quantifying the relationship between the risk and the explanatory factors, which in our context we may consider to be “locational”. The simplest quantification of a locational factor is distance from F and we will be very largely concerned with the analysis of distances in what follows. It is important to realise,

however, that distance from a point is by no means the only locational measure; in principle we might consider a variety of geographical variables such as distance from a coastline, altitude, angular deviation from the direction of the prevailing wind, or any function of such quantities. What we choose will depend on our idea of aetiological mechanisms (or, in statistical terms, of alternative hypotheses). Some investigators object on the grounds that a postulated association with simple distance from F greatly oversimplifies the geographical aspects of the aetiology. We would argue, however, that no analysis of location is possible without some underlying concept of spatial relationship and a quantification of it, by distance or otherwise. At the same time, it is recognised that crude Euclidean distance is certainly not appropriate as a quantitative variable for calculation of associations, because area increases as the square of distance and so more weight is given to the most distant subregions. This is the opposite of what we want, so it will be useful to think of a suitable inverse measure, which we may call “closeness”. The choice of transformation we use to achieve this introduces an element of arbitrariness. It will perhaps be preferable to use methods that are invariant under monotonic transformations—ie, are based on ranks of distances—while at the same time according more weight to points that are close to F than to those that are distant.

Poisson-distributed counts with an implicit ordering

A GENERAL TEST

These considerations lead naturally to a class of methods explored in detail by Stone.⁶ He considers N regions around F for which we can calculate the expected numbers E_i of cases of the disease under consideration, using appropriate national or regional rates. The regions are supposed to be ordered with respect to distance from F. We may then consider that the corresponding counts C_i are independently Poisson-distributed with means $\lambda_i E_i$, where the null hypothesis, that F introduces no extra risk, takes the form

$$H_0: \lambda_1 = \lambda_2 = \dots = \lambda_N = 1.$$

A very general form of alternative hypothesis is

$$H_1: \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N,$$

with at least one inequality holding, and a test powerful against this whole class of alternatives would not be dependent on the scale of distance implied by the aetiological mechanism. We would then assume only that the risk would never increase with increasing

distance from F; violation of this condition would not invalidate such a test, though the statistical power may well be reduced.

A general principle of the theory of significance testing is to compare the likelihood of the data L_1 under H_1 with L_0 , the corresponding likelihood under H_0 , choosing any parameter estimates to maximise L in either case. The resulting maximum likelihood ratio (MLR) statistic generally has good properties when testing H_0 against H_1 . In this particular case the maximum likelihood estimates under H_1 can be calculated by a method related to the theory of isotonic regression,¹⁷ though the method generally requires the use of a computer. The distribution of the resulting MLR statistic is correspondingly difficult to derive and in practice it would be necessary to use simulation methods to estimate the significance level of a particular result.

THE POISSON MAXIMUM TEST

It turns out that, in many cases of interest, the estimator $\hat{\lambda}_1$ of λ_1 under H_1 contains much of the information available for testing H_0 .

It is comparatively easy to prove that

$$\hat{\lambda}_1 = \max_{1 \leq n \leq N} \left[\frac{\Sigma C_i}{\Sigma E_i} \right]$$

ie, it is the largest empirical relative risk (or SMR) that we would observe if we calculated all the ratios for areas obtained cumulatively by including successive sub-regions. This simple interpretation invests this "Poisson maximum" statistic with considerable appeal; it provides an adaptive test, in the sense that the data will select the distance at which the observed effect is maximal. It is of course necessary to allow for this element of selection when evaluating the distribution of $\hat{\lambda}_1$ under H_0 , but fortunately this distribution can be calculated exactly by using methods related to the theory of random walks. Again a computer is necessary in general, although in some circumstances simple approximations to the significance level may be calculated. See Stone⁶ for details and for power comparisons which confirm that this Poisson maximum test generally compares favourably with the MLR test against general ordered alternatives.

APPLICATION TO SELLAFIELD DATA

The above methods have been applied to the 36 cases of children registered as having leukaemia, diagnosed before the age of 15 between 1968 and 1982 in Cumbria.^{18 19} The data were obtained from the

Northern Region Cancer Registry and were kindly provided by Dr Craft of the University of Newcastle upon Tyne, who obtained from the postcodes grid references accurate to 100 metres. For the purposes of these analyses the 36 children were allocated to their electoral wards for which age-sex adjusted expectations were also calculated, using OPCS small area statistics.

The remarkable feature of these data is well known, namely that Seascale, the nearest ward to Sellafield, experienced four cases of leukaemia in these ranges of age and time, against an expectation of only 0.196. The significance level of this single observation, obtained from the cumulative probabilities of the Poisson distribution, is 5.245×10^{-5} .

Figure 1 shows the cumulative observed and expected numbers of cases near Sellafield and the

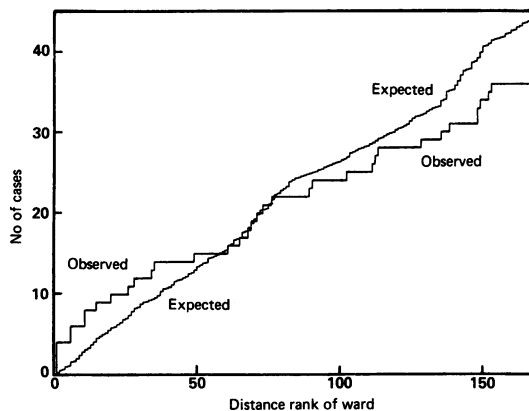


Fig 1 Cumulative observed and expected numbers of cases in electoral wards ranked by distance from Sellafield.

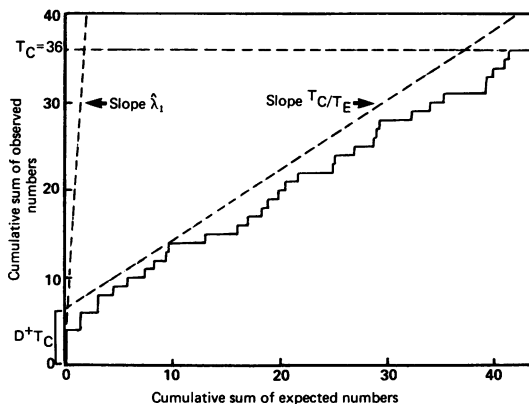


Fig 2 Cumulative sum of observed numbers plotted against cumulative sum of expected numbers. Sellafield data on childhood leukaemia.

initial step corresponding to the four cases in Seascale can be seen clearly. An alternative way of displaying the data is to plot ΣC_i against ΣE_i (fig 2) which displays the Poisson maximum statistic as the slope of the least steep line through the origin that lies above the graph. This is clearly determined by the first observation, ie, Seascale itself, and it is not too surprising that the significance level for this statistic, $p = 5.24 \times 10^{-5}$, is almost the same as the selected value for Seascale only. This is essentially because of the very low probability that the locus of this graph should subtend a greater angle than this further out. The test is clearly highly influenced by an early excess; the fact that in the region as a whole the total number observed, T_C , say (= 36), is less than that expected, T_E , (= 44.12) does not weaken the test in its power to detect the early excess. It should be emphasised, however, that the test adapts itself to a more attenuated gradient of excess risk and that in all cases the significance level calculated allows for the selection involved.

OTHER METHODS

The Poisson maximum and MLR tests provide at least a partial answer to the difficulties described in (i)-(iv) above. They still require knowledge of population sizes in small regions, ascertainment of all relevant cases and suitable comparative rates for calculating expectations. If it is thought that the expectations calculated are consistently too high or too low, a *conditional* analysis may be performed. This is executed by multiplying the expectations by T_C/T_E , so that now the adjusted expectations add to T_C . The MLR and Poisson maximum statistics can be calculated as before, but generally it is necessary to estimate the significance level by simulation. This is conceptually quite simple: T_C cases are allocated at random to the N subregions according to probabilities E_i/T_C and the test statistic is calculated for this sample; the process is repeated a large number of times and the actual value of the statistic is compared with the resulting distribution.

In effect the adjustment of the expectations forces the two cumulative graphs of fig 1 to go through the same terminal point. If we plot both on a scale of relative frequency vertically and replace the distance rank by the distance itself on the horizontal axis (fig 3) it is immediately clear that we are comparing the distribution of the distances of the cases with an expected or theoretical distribution calculated from the population rates, which may of course be age-sex adjusted.

This comparison may in principle be achieved by any of a large number of tests, but, as remarked above, absence of any clear idea about the quantitative connection between distance and risk means that we shall be well advised to restrict ourselves to tests based

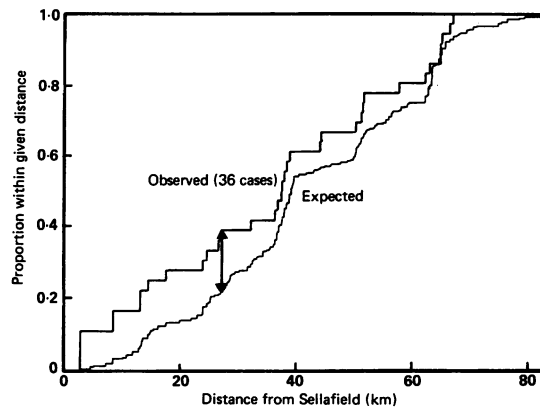


Fig 3 Cumulative distributions of observed and expected distances of childhood leukaemia cases from Sellafield. The vertical line indicates D^+ , the maximum deviation between the distributions.

on distance ranks. The Kolmogorov-Smirnov test is a good general purpose test which will be sensitive to differences in risk at all distances, but with no specific sensitivity in the tails of the distribution. The test is based on the largest vertical difference between the two distributions, which depends only on the sample size and not on the underlying distribution being tested, provided the latter is continuous. In our case the distribution is discrete, with $N = 168$ points, and it is desirable to make an adjustment, as described by Conover.²⁰ When this is done the statistic for the one sided test of the Sellafield data is $D^+ = 0.171$, resulting from a maximum deviation at $n = 35$. This achieves significance only at the 10% level. Figure 2 demonstrates the connection between the Poisson maximum and the Kolmogorov-Smirnov tests, showing how the latter represents the smallest intercept of a line having slope T_C/T_E and lying entirely above the locus of cumulative counts.

It is worth pointing out that, by regarding the cases as defining a distribution of distances, we have moved from a model in which the random observation is disease occurrence (in fixed locations) to one in which the variable is distance and the cases are regarded as fixed. This is precisely the duality that exists between prospective ("cohort") studies and retrospective ("case-control") studies. The transition is justified by the same kinds of argument and offers interesting possibilities for further methods. Amongst these is the use of a suitable *sample* of controls to estimate the distribution of the distances of members of the population from F; this will be the subject of a further paper.

Application to Sizewell data

The methods using population based expectations have also been applied to cases of leukaemia registered at all ages between 1967 and 1981 in East Suffolk in relation to their distances from the nuclear power station at Sizewell. The analysis included just the 48 parishes within 17 km of Sizewell, this being the largest distance contained within the area covered by the report from which the data came, published by the District Medical Officer for East Suffolk.²¹ The expectations in the report were based on crude population totals without correcting for differences in the age-sex distribution between parishes. This could be important because of the strong increase in leukaemia incidence with age. The expectations in the present analysis were therefore recalculated using age and sex specific population sizes and the specific rates obtained from the East Anglian Cancer Registry for the same time period. The census figures for 1971 were used as this was the available date nearest to the centre of the period considered. Finally distances were

Table *Observed and expected numbers of leukaemia cases registered in the Sizewell region (all ages).*

Rank	Observed	Expected	Parish
1	10.000	7.00100	Leiston
2	1.000	1.46800	Aldringham
3	0.000	0.58740	Theberton
4	1.000	0.85110	Knodishall
5	0.000	0.65810	Middleton
6	0.000	0.25230	Dunwick
7	10.000	4.42000	Aldeburgh
8	1.000	0.97400	Friston
9	0.000	0.95230	Westleton
10	1.000	0.29620	Sternfield
11	0.000	0.58730	Darsham
12	6.000	2.64700	Saxmundham
13	1.000	0.16430	Iken
14	1.000	0.86340	Snape
15	1.000	1.60700	Kelsale
16	1.000	1.11600	Yoxford
17	0.000	0.72670	Benhall
18	0.000	0.19330	Farnham
19	0.000	0.08959	Thorington
20	0.000	1.43600	Blythburgh
21	0.000	0.38100	Sibton
22	0.000	0.96780	Walberswick
23	0.000	0.29020	Stratford St Andrew
24	0.000	0.45240	Rendham
25	0.000	0.27190	Sweffling
26	0.000	0.68490	Bramfield
27	1.000	0.50970	Sudbourne
28	1.000	0.57240	Blaxhall
29	1.000	0.67540	Tunstall
30	0.000	0.23090	Chillesford
31	0.000	1.55700	Wenhaston
32	0.000	0.34580	Little Glenham
33	0.000	0.31920	Great Glenham
34	0.000	1.29600	Orford
35	0.000	0.88110	Peasanhall
36	0.000	0.30060	Walpole
37	0.000	0.34640	Marlesford
38	0.000	0.20200	Bruisyard
39	0.000	0.56980	Rendlesham
40	0.000	0.20770	Cransford
41	0.000	0.08114	Wantsiden
42	0.000	0.44980	Camsey Ash
43	0.000	0.15960	Haveringham
44	0.000	0.47270	Parham
45	0.000	0.32620	Butley
46	0.000	0.06584	Gedgrave
47	0.000	0.51180	Hacheston
48	0.000	0.64840	Badingham

measured using the centroids of the parish populations calculated as the arithmetic means of the grid coordinates for the separate enumeration districts and published by OPCS.

The data are shown in the table and it will be seen that again the nearest parish (Leiston) shows an excess of cases observed over expectation. The Poisson maximum statistic, however, occurs at $n = 13$ (Iken) largely because of the extra excess at Aldeburgh ($n = 17$). The value of the statistic is 1.486, yielding a significance level of $p = 0.16$, which is not significant at any level of interest. The one sided Kolmogorov-Smirnov statistic, however, is 0.331 (with the maximum discrepancy at $n = 14$, Snape) and this gives a significance level of less than 0.005 after adjusting for the discreteness of the data.

Bush²¹ discusses possible explanations for a raised incidence, though his own analysis, not being formally based on distance, did not demonstrate a significant excess. He points out that three of the Leiston cases were power workers, but data on the residence locations of power workers as a whole are not available, so we are unable to investigate this factor. Stone⁶ presents further analyses and examines the change of population structure in Suffolk in the period considered.

Discussion

We have in this paper discussed aspects of tests for locational data from two principal standpoints: that they should be good at detecting cases near a source of risk F and that they should be sensitive to the influence of individual cases which might otherwise be lost in the large numbers typical of registry data.

In particular, the "Poisson maximum" approach is clearly very sensitive to cases near to F and may be used with advantage where the excess risk is concentrated there. It has the attraction of having a ready interpretation as a maximum observed risk ratio, with due allowance for selecting this maximum over all possible distances. For such situations—as in the Sellafeld case—it is virtually immaterial how large a region is studied altogether. This is in contrast to methods which merely examine incidence rates. The fact that the test may lose power against a non-monotonic effect is illustrated by the non-significant result applying to the Sizewell data. It goes without saying that for a formal test of significance we should decide what test we wish to do in advance.

The emphasis in this paper has been on methodology, especially in relation to the need for methods that will appear to be fair and convincing to those who are impressed by small case series. The data sets themselves are clearly of great interest and deserve at least some tentative conclusion. It seems clear that

the excess observed at Seascale is not merely an artefact due to data selection by region. The significance level observed makes it hard to believe that the cases are coincidental and we are inclined to believe that the excess results from some genuine, but as yet unidentified, cause. At the same time there is very little evidence of any other locational association once the Seascale cases have been removed. Recent discussions of the data relating to Sellafield include that of Hargreaves *et al.*²² The Sizewell data are altogether less statistically significant. Geographical proximity to the power station does not seem to be particularly important; the possibility of an occupational effect is open but seems unlikely in view of the careful monitoring carried out in the industry.

We wish to thank Dr A W Craft of the University of Newcastle upon Tyne for kindly making the Cumbrian data available. Figure 2 is reproduced from *Statistics in medicine*⁷ by permission of John Wiley and Sons Ltd.

Address for correspondence and reprints: Dr J F Bithell, Department of Statistics, 1 South Parks Road, Oxford OX1 3TG.

References

- ¹ Preston DL, Kato H, Kopecky KJ, Fujita S. *Life span Study Report 10: Part 1, Cancer mortality among A-bomb survivors in Hiroshima and Nagasaki, 1950-82*. Hiroshima: Radiation Effects Research Foundation, 1986. RERF Technical Report TR 1-86.
- ² Bithell JF, Stewart AM. Pre-natal irradiation and childhood malignancy: a review of British data from the Oxford Survey. *Br J Cancer* 1975; **31**: 271-87.
- ³ Kneale GW, Stewart AM. Prenatal X-rays and cancers: Further tests of data from the Oxford Survey. *Health Phys* 1986; **51**: 369-76.
- ⁴ Bithell JF, Stiller CA. A new calculation of the carcinogenic risk of obstetric X-raying. *Stat Med* 1988; **7**: 857-64.
- ⁵ Yorkshire Television. *Windscale, the nuclear laundry*. Documentary programme, 1983 Nov 1.
- ⁶ Stone RA. Statistical methodology and causal inference in studies of the health effects of radiation. Oxford: University of Oxford, 1986. 208 pp. DPhil thesis.
- ⁷ Stone RA. Investigations of excess environmental risks around putative sources: statistical problems and a proposed test. *Stat Med* 1988; **7**: 649-60.
- ⁸ Knox, EG. The detection of space-time interactions. *Appl Stat* 1964; **13**: 25-9.
- ⁹ Smith PG. Spatial and temporal clustering. In: Schottenfeld D, Fraumeni JF, eds. *Cancer epidemiology and prevention*. Philadelphia: WB Saunders, 1982: 391-407.
- ¹⁰ Barton DE, David FN, Fix E, Merrington M, Mustacchi P. Tests for space-time interaction and a power function. In: LeCam LH, Neyman J, eds. *Proceedings of the fifth Berkeley Symposium*. Berkeley, California: University of California Press, 1967; **IV**: 217-28.
- ¹¹ Bradshaw D. A comparison of tests for space-time clustering of disease cases. Oxford: University of Oxford, 1982. 123 pp. DPhil thesis.
- ¹² Craft AW, Birch JM. Childhood Cancer in Cumbria. *Lancet* 1983; **ii**: 1299.
- ¹³ Gardner MJ, Winter PD. Mortality in Cumberland during 1959-78 with reference to cancer in young people around Windscale. *Lancet* 1984; **i**: 216-7.
- ¹⁴ Urquhart J, Palmer M, Cutler J. Cancer in Cumbria: the Windscale connection. *Lancet* 1984; **i**: 172-3.
- ¹⁵ Wakeford R. Radiological Protection Bulletin. 1987: 83.
- ¹⁶ Craft AW, Openshaw S, Birch JM. Childhood cancer in the Northern Region, 1968-82: incidence in small geographical areas. *J Epidemiol Community Health* 1985; **39**: 53-7.
- ¹⁷ Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD. *Statistical inference under order restrictions: the theory and application of isotonic regression*. London: John Wiley, 1972.
- ¹⁸ Black D. *Investigation of the possible increased incidence of cancer in West Cumbria*. London: HMSO, 1984.
- ¹⁹ Craft AW, Openshaw S, Birch J. Apparent clusters of childhood lymphoid malignancy in Northern England. *Lancet* 1984; **ii**: 96-7.
- ²⁰ Conover WJ. A Kolmogorov goodness-of-fit test for discontinuous distributions. *J Am Stat Assoc* 1972; **67**: 591-5.
- ²¹ Bush MFH. *Leukaemia in East Suffolk*. Unpublished report by East Suffolk Health Authority, 1983.
- ²² Hargreaves R, Wilkie D, Wakeford R. A re-examination of the epidemiological data for Cumbrian coastal areas. In: *Health effects of low dose ionising radiation—recent advances and their implications*. Proceedings of a conference of the British Nuclear Energy Society and the European Nuclear Society. London: British Nuclear Energy Society, 1987.