

Digesting the Seascale data

SIR — The report of the Independent Advisory Group on the possible increased incidence of cancer in West Cumbria¹ has raised issues about the interpretation of occurrence rates for lymphoid malignancy, and of the associated cumulative Poisson probabilities, in the electoral wards of the Northern region.

M.J. Gardner (*Nature* 312, 16; 1984) rightly suggested that these two 'indices' bear upon different aspects of the data. The occurrence rate, which relates the recorded number of malignancies (for a particular ward over a defined period of time) to the size of the population at risk, may not be an accurate reflection of the 'true' or 'underlying' incidence rate because of the small size of the at-risk population involved in most of the wards. The recorded number of malignancies over any finite period of time is subject to an element of 'random' influence, taken account of in cumulative Poisson probabilities given in the Advisory Group report and by A. Pomiankowski (*Nature* 311, 100; 1984) which express the chance of getting the observed occurrence rate, or a higher one, if the underlying incidence rate for the ward in question is equal to the regional average rate of 0.61 per 1,000.

These Poisson probabilities, like the observed occurrence rates, must be interpreted in the light of the ward population size. The same observed occurrence rate of 5 per 1,000 in wards with at-risk populations of 200 and 400, for example, would give cumulative Poisson probabilities of 0.115 and 0.025 respectively. In other words, as Gardner pointed out, the Poisson probabilities are not a direct indicator of occurrence rate.

An alternative way of presenting the figures is by constructing a confidence interval for the underlying incidence rate. Confidence intervals at 95% and 99% confidence levels, for the wards with six highest observed rates of lymphoid malignancy, are shown in the table below. Although this alternative mode of presentation does not lend itself quite so readily to ranking as occurrence rates and Poisson probabilities, it does bring out two points with some clarity. First, the width of the intervals is a measure of how

relatively imprecise inferences made from these data are, because of the small ward sizes involved. The observed occurrence rate of 9.7 per 1,000 in the Seascale ward, for example, is consonant, at 95% confident level, with a true rate of anything between 2.65 and 24.9 per 1,000. Second though, despite this imprecision, the intervals give an indication of why the figures for Seascale are a cause for concern. Even at the higher confidence level of 99%, the minimum estimate of true incidence rate consonant with the data is 1.64 per 1,000, more than two-and-a-half times the regional average occurrence rate.

A related aspect of the small ward sizes, and the consequent imprecision of inferences made from the data, is the difficulty of detecting when the underlying incidence in a small area is raised above the regional average, even by substantial amounts. R. Wakeford (*Nature* 312, 191; 1984) expressed agreement with the Advisory Group conclusion that the Poisson probability calculated for the Seascale ward (0.000134), though extreme, does not justify the conclusion that Seascale is 'unique'. One must bear in mind, I think, that even if the true incidence rate in the Seascale ward were, say, five times the regional average, the chance of recording five or more occurrences (and thereby getting Poisson probability more extreme than 0.000134) would be less than 1%. The odds are clearly stacked against being able to reach any firm conclusions on occurrence data alone, even if a seriously inflated true incidence rate pertains.

Seascale was the only ward in western Cumbria picked out by Craft, Openshaw and Birch³ as having a Poisson probability less than 0.05. R. Wakeford suggested that this observation does not support the hypothesis of causal relationship between excess childhood lymphoid malignancies and increased radiation exposure around the Sellafield site. However, the small ward sizes, together with the 'discrete' nature of the Poisson probability distribution, imply that some caution must be exercised in interpreting a geographical pattern derived by reference to a specific cut-off value: in this case, the conventional 5% significance level. Two examples should suffice

to illustrate the point. Consider two hypothetical wards, one with a true incidence rate three times the regional average and an at-risk population size 100, the other with twice the regional average incidence rate and an at-risk population of 250. To get a Poisson probability of less than 0.05 in the first ward, 2 or more occurrences would need to be recorded (giving a Poisson probability of 0.0018, a good deal less than 0.05). The chance of this happening is less than 1.5%, even though the true incidence rate in this hypothetical ward is three times the regional average. In the second ward, again 2 or more occurrences would need to be recorded (giving a Poisson probability of 0.0105). The chance of this happening is less than 4%, despite the assumption of a true incidence rate double the regional average.

EDDIE SHOESMITH

University of Buckingham,
Buckingham MK8 1EG, UK

1. Report of the Independent Advisory Group, Chairman Sir Douglas Black, *Investigation of the Possible Increased Incidence of Cancer in West Cumbria* (HMSO, London, 1984).
2. Pearson, E.S. & Hartley, H.O. *Biometrika Tables for Statisticians*, Vol. 1, 3rd edition (Cambridge University Press for the Biometrika Trustees, 1970).
3. Craft, A.W., Openshaw, S. & Birch, J. *Lancet* ii 96-97 (1984).

Pinning down stem-cell populations

SIR — The proposal¹ that the only valid representatives of multipotent, 'persistent' haematopoietic stem cells are 'late' (≥ 12 day) spleen colonies has been critically discussed by Blackett². In support of such a critique, I quote a 1964 report³ showing that a number of 'early' (10-day) erythroid, unipotent and, according to ref. 1, 'transient' colonies do contain multipotent stem cells. In addition it has been shown that multipotent stem cells within many 'persistent', late (12-, 13-day) colonies do have transient potential for self renewal⁴. To avoid circular argument I propose that 'early' spleen colony counts adequately estimate the stem cell contents in a given sample. These counts are the most appropriate because they minimize the time-dependent colonial increase by stem cells migrating into test spleens from distant locations⁵. Obviously, the differentiative potentialities and renewal efficiency of stem cells within such colonies cannot be inferred from direct counts. These parameters have to be indirectly estimated by secondary transplantation procedures and interpreted exclusively in such contexts^{2,3,5}.

G.T. MATIOLI

University of Southern California
School of Medicine,
2025 Zonal Avenue,
Los Angeles, California 90033, USA

1. Magli, M.C., Iscove, N.N. & Odarchenko, N. *Nature* 295, 527 (1982).
2. Blackett, N.M. *Nature* 302, 271 (1983).
3. Lewis, J.P. & Trobaugh, F.E. *Nature* 204, 589 (1964).
4. Vogel, H., Niewisch, M. and Matioli, G. *J. cell. Physiol.* 72, 221 (1968).
5. Matioli, G.J. *reticuloendothel. Soc.* 15, 282, 1974.

West Cumbria lymphoid malignancy cases

Cases	Children	Observed rate per 1,000	Confidence intervals for true incidence rate	
			(95%)	(99%)
2	144	13.9	[1.68, 50.1]	[0.72, 64.4]
1	97	10.3	[0.26, 57.4]	[0.05, 76.6]
4*	411	9.7	[2.65, 24.9]	[1.64, 30.6]
1	165	6.1	[0.15, 33.8]	[0.03, 45.0]
1	172	5.8	[0.15, 32.4]	[0.03, 43.2]
1	174	5.7	[0.15, 32.0]	[0.03, 42.7]

Based on Table 40 of ref 2, showing 95% and 99% confidence levels.

* Seascale.