
Reply

Lawson's letter makes a number of points concerning the importance of spatial autocorrelation in disease data with which I cannot agree. This is an area where generalisation may be dangerous, so I would certainly not wish to dispute the possibility that such autocorrelation may be present, in some instances. But by the same token, I believe that there are others where it is not of major concern. I restrict my remarks to diseases for which cases may be taken to be independent of one another conditionally on any underlying variation in risk; the appropriate underlying error model must then be the Poisson process, although we will not necessarily know the true risk at a given point.

In the first place, unrecognised spatial autocorrelation presumably has the effect – at least within the framework of frequentist inference and modelling – of leading to heterogeneity that might be recognised, for example, by a high deviance residual to a fitted model. However, in this respect this heterogeneity is like that resulting from any other factor that has not been properly taken into account in a given model. The crucial question is surely whether this over dispersion leads to misleading conclusions about the factors that have been fitted. Whether it does so will, broadly speaking, depend on whether these factors are themselves associated with whatever spatially varying factors are responsible for the over dispersion. If they are, then of course such factors may be regarded as confounding, since we are interested in whether these factors are themselves associated with whatever these varying factors are responsible for the over dispersion.

However, as 1) above and 2) below, we can never eliminate the possibility of confounding factors anyway; it seems to me to be of limited inferential value to know that such confounding – if it occurs – is mediated through an unidentified spatially varying variable. It is, of course, scientifically much more valuable to try and identify all the factors that matter.

Lawson claims that the heterogeneity introduced by spatial autocorrelation is "particularly important in putative source applications". I believe this to be most unlikely. For here the contributions of cases to neighbouring intervals on the distance axis will come in large measure from small areas which are relatively far apart, so that the serial correlation on this axis should be expected to be substantially less than the spatial autocorrelation per unit distance prevailing in two dimensions.

Likewise, the claim that "Leukaemias are known to cluster..." and putative sources (nuclear power stations) is highly suspect.

There is very little published evidence relating to adult leukaemias, while, for childhood leukaemia in England and Wales, the evidence was recently reviewed in the *British Medical Journal* and found to be largely negative, with the exception of the data relating to Sellafield. The latter excess is extremely hard to interpret because of the post hoc nature of the observation; numerous papers have addressed the possible aetiology, but it seems clear that a purely geographical explanation is far from convincing.

In fact, even if the numerous claims in the literature, childhood leukaemia does not exhibit strong spatial-auto-correlation. The residual deviance exhibited in my own contribution to the session addressed in Lawson's letter suggests that there is little residual heterogeneity to explain, while methodological sound and intelligible attempts to demonstrate a spatial explanation for any such heterogeneity have generally produced only equivocal results.

For adults, there tends to be substantially more heterogeneity, but the small scale spatial component of this again appears to be very weak.

To assume that spatial auto-correlation is important and to build this assumption into a statistical analysis may at first sight appear to be a sensible defensive strategy. But there is a high price to pay for an unnecessarily elaborate model. This situation becomes more difficult and it is likely that estimates of the parameters of primary interest become less precise and stable as we attempt to gain more information from modest data sets. Important questions of model sensitivity or "parsimony" are also much harder to address. These are familiar ideas in applied statistics; the fact that they are hard to make precise in any given case does not make them less important.

The onus is on the builder of a more complex model to demonstrate that a simpler one will not suffice.

J BITHELL
Department of Statistics
University of Oxford


The decline in sex ratios at birth, England and Wales, 1973-90

Sin – Dickinson and Parker entitled the letter which was published in this journal, "Why is the sex ratio falling in England and Wales?". The short answer to this question is that the sex ratio of newborns is falling still. It fell more or less continuously from 1973-90, but in the four succeeding years it rose; and in 1993 it reached the highest level since 1983. It is interesting, however, to consider the causes of the fall during the 1970s and 1980s. Dickinson and Parker mention three possibilities – coital rate, pollution, and hormones – and I should like to comment on these.

Coital rate

The sex ratios of white births in the USA between 1915 and 1988 seemed to show unexplained oscillations, and there were two very roughly 20-30 years. In particular, they rose during the years 1965-75 and fell during the years 1975-88. In that paper, 1 cited evidence that age standardised reported marital coital rates increased by 22% between 1965 and 1975 and fell by 27% between 1975 and 1988. Moreover, I also reviewed the hypothesis that parental coital rate (weakly) controls the offspring sex ratio. 2 The evidence seems strong so it appears that in the USA from 1965-88 variations in the sex ratio might be explained by secular variations in the coital rate.

The recent decline in the sex ratio in England and Wales parallels a decline in the USA across the same years. It is possible that the

Reply

Lawson raises a number of interesting issues, but in our opinion misrepresents what we say in our paper in several respects.

1 We agree that the assumption of an underlying Poisson point process may be violated in practice, which is why we refer to this model as an idealised point process model. We discuss this issue explicitly in the last two paragraphs of our paper.

2 We agree that residential location is an imperfect surrogate for the epidemiologically relevant location or locations of an individual, but again we say this explicitly in the paper. Also, dealing with counts in regions hides, rather than resolves, the difficulties which arise from introducing other potential problems, most obviously the well documented one of ecological bias. Nowhere in our paper do we suggest that our approach (or any other single approach) is a "general panacea"; we do try to clarify some of the relevant practical issues by deliberately working within an idealised theoretical framework.

3 The problem of "mixed aggregate levels" is much more subtle than simply being able to deal with point cases and aggregated population data (controls). In our opinion, much work remains to be done in developing practical solutions to this problem.

4 Models more complex than our simple equation (8) would undoubtedly be justified if sufficient, and sufficiently precise, data were available for their validation. It does not follow that such models would give better inferences for the necessarily limited and imprecise data available in many applications. We have proposed a general model for decline in risk with distance around the source that approximates to the "threshold" or "point exposure" (eg around a stack) scenarios discussed by Lawson (Diggie et al, unpublished data).

While recent methodological developments in complex statistical modelling open up exciting possibilities, it should not be forgotten that all models are wrong, and that parsimony is a sound guiding principle.

P DIGGLE
Department of Mathematics and Statistics,
University of Lancaster
P ELJJOTT
London
Hormones

The proposition nevertheless arises whether the decline in sex ratios in England and Wales from 1973–90 reflected some sort of increasing hazardous environmental exposure to men. In particular, is it to be associated with the recent suggestion relating diminished sperm counts to environmental oestrogen exposure? 14 If the USA is anything to go by, this seems not to be the case. There (where the data on secular movements in sperm counts are more abundant than in the UK) sperm counts were apparently declining during the 1970s and possibly stable during the 1980s, 15 in contrast to the movements of the sex ratio described above.

David G. Taylor 1

REFERENCES


