

DEALING WITH CONFOUNDING WHEN INVESTIGATING TIME-SPACE CLUSTERING OF ANIMAL DISEASE

Ward MP

Department of Primary Industries, Animal Research Institute, Locked Mail Bag 4,
Moorooka, Queensland, Australia 4105

Present address: Department of Veterinary Pathobiology, School of Veterinary
Medicine, Purdue University, West Lafayette, Indiana, U.S.A. 47907

A one-dimensional scan statistic has been used to study disease clusters in time,¹ and a two-dimensional scan statistic has been proposed for spatial disease clusters.² Kulldorf et al³ have proposed extending the use of the scan statistic, defined by a cylindrical window with a circular geographic base and height corresponding to time, for time-space clustering. By assuming the number of disease cases within the scanning window to be Poisson distributed, confounders can be controlled through indirect standardisation. This statistic also accommodates the uneven distribution of most populations. Multiple testing can be resolved by Monte Carlo simulation.

Blowfly strike is a major economic disease in Australian sheep flocks.⁴ Breech and body strike are the most important types of flystrike.^{5,6} Risk factors for breech strike include diarrhoea and urine soiling, and weaner sheep and ewes are therefore more susceptible.⁵⁻⁷ Management procedures (eg. mulesing, crutching, tail docking, intestinal parasites control) are used to reduce susceptibility.^{6,7} Body strike appears to be strongly determined by climatic factors.⁶ Body and breech strike data were used in this study to assess application of the scan statistic to detect time-space clustering whilst controlling the potential confounding of flock structure.

Materials & Methods

The following information was obtained from a survey of 32 producers located within two adjacent local government areas of southeastern Queensland: flock size and structure (number of lambs, wethers, ewes and rams) at the most recent shearing and number and type of flystrike cases occurring between August 1998 and May 1999.⁸ Clustering of body strike and of breech strike was analysed using the scan statistic and a Poisson model.³ Scanning windows were limited to half the expected cases; clusters with either high or low rates of strike were included. Likelihood ratios were used to test statistical significance of clusters, their distribution under the null-hypothesis and P-value being obtained by Monte Carlo simulation ($\alpha=0.05$). Secondary clusters were also identified by their likelihood ratio and whether they overlapped the primary cluster. Crude analyses were re-analysed adjusting for flock structure. All calculations were performed using SaTScan software version 2.1 (<http://dcp.nci.nih.gov/BB/SaTScan.html>).⁹

Results

Significant ($P < 0.01$) clustering of body strike was detected in crude and adjusted analyses (log likelihood ratio 923 and 945, respectively) and the primary cluster was identified at the same location during March 1999 (relative risk 3.73 and 3.99, respectively). Significant ($P < 0.01$) clustering of breech was detected between February and May 1999 in crude and adjusted analyses (log likelihood ratio 541 and 475, respectively), but the primary cluster identified consisted of an additional 3 flock when flock structure was controlled (Figure 1). The relative risk of this cluster (6.83) changed substantially when controlled for flock structure (relative risk 4.57).

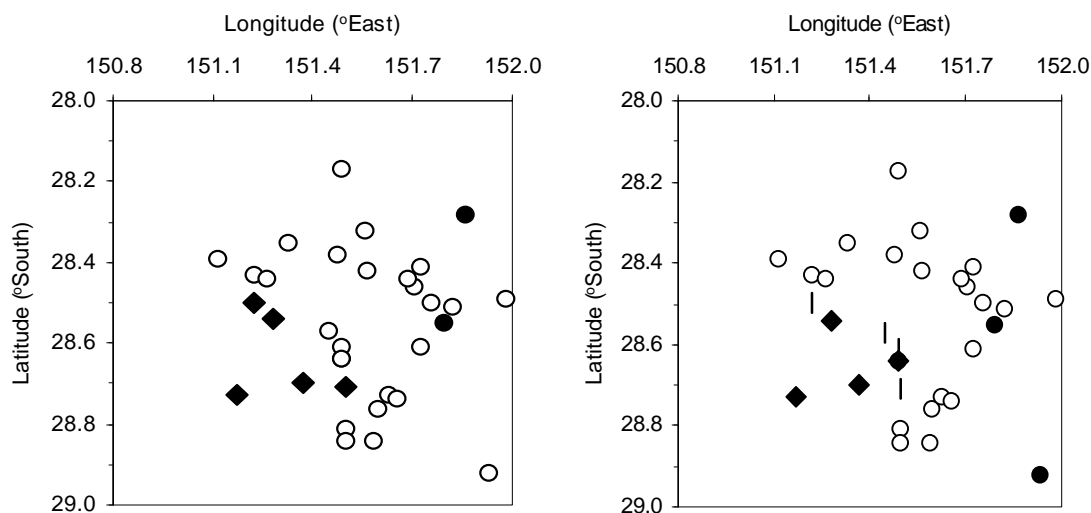


Figure 1. Crude (left) and flock structure adjusted clustering of breech strike in southeastern Queensland sheep flocks (u primary cluster; l secondary cluster).

Discussion

Investigation of time-space clustering of disease is potentially a powerful method of generating and testing hypotheses of disease causation. To fully examine time-space clustering objectively, the use of statistical tests is necessary.¹⁰ Global, focused and detection tests are available. Global procedures, such as Knox, Mantel and nearest neighbour tests and Barton's method, test for clustering throughout the study area, but do not identify specific clusters. Focused methods are used when there is a prespecified point source. In contrast, cluster detection methods such as the scan statistic are able to identify and test the significance of specific clusters. The scan statistic can also adjust for the uneven population density present in almost all populations, analysis being conditioned on the total number of cases observed. In addition, cluster definition is flexible: it may be individuals or aggregated data (eg. herds, local government areas, postal codes, states). Use of the Poisson model allows control of any number of covariates, using indirect standardization.

Using the scan statistic, clustering of both body strike and breech strike was detected in the study area between August 1998 and May 1999. Results suggest that the structure of flocks needs to be taken into account when investigating breech strike: if flocks contain a greater proportion of weaner sheep and/or ewes, more cases of breech strike may be expected. Confounding identified in this study was substantial: estimated relative risk decreased by 33% after adjusting for flock structure. Failure to adjust for flock structure may therefore result in spurious clustering of disease being identified.

Few examples of investigating the interaction of time and space in disease clustering exist in the veterinary literature.¹⁰ Reasons may include inappropriate study design, difficulty visualising and explaining the concept of time-space interaction, the implicit assumption of homogeneity of the population of interest in some techniques, lack of accessibility to suitable analytical software, and publication bias due to insufficient power of available analytical methods. Development of monitoring and surveillance systems in veterinary epidemiology are likely to increase the need for time-space analyses. The scan statistic appears to be a suitable technique for analysing such data sets for clustering. However, to fully explore such data sets whilst controlling confounding, enumeration of the population of interest is necessary. Collecting data on the population of interest and a range of potentially confounding factors will enable the greatest amount of information to be gained from monitoring and surveillance systems.

References

- 1 Naus J. The distribution of the size of maximum cluster of points on the line. *Journal of the American Statistical Association* 1965; 60: 532-538.
- 2 Kulldorff M. A spatial scan statistic. *Communications in Statistics: Theory and Methods* 1997; 26: 1481-1496.
- 3 Kulldorff M, Athas WF, Feuer EJ, Miller BA, Key CR. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos. *American Journal of Public Health* 1998; 88: 1377-1380.
- 4 McLeod RS. Costs of major parasites to the Australian livestock industries. *International Journal of Parasitology* 1995; 25: 1363-1367.
- 5 Wardhaugh KG, Bedo D, Vogt WG. Management model for the control of flystrike on sheep. In: *Proceedings of the 1994 Australian Sheep Veterinarian Society meeting, Canberra, 70-73.*
- 6 Wardhaugh KG, Morton R. The incidence of flystrike in sheep in relation to weather conditions, sheep husbandry, and the abundance of the Australian sheep blowfly, *Lucilia cuprina* (Wiedemann) (Diptera: Calliphoridae). *Australian Journal of Agricultural Research* 1990; 41: 1155-1167.
- 7 Watts JE, Murray MD, Graham NPH. The blowfly strike problem of sheep in New South Wales. *Australian Veterinary Journal* 1979; 55: 325-334.
- 8 Ward MP, Carpenter TE. Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. *Preventive Veterinary Medicine* 2000; in press.
- 9 Kulldorff M, Rand K, Gherman G, Williams G, DeFrancesco D. SaTScan v2.1: Software for the spatial and space-time scan statistics. Bethesda, MD: National Cancer Institute, 1998.
- 10 Ward MP, Carpenter TE. Analysis of time-space clustering in veterinary epidemiology. *Preventive Veterinary Medicine* 2000; 43: 225-237.