Comment on historical article

Secular pattern of congenital oesophageal atresia—George Knox, 1959

Helen Dolk

This paper on the clustering of a congenital malformation was an early and important contribution to an issue which is still very much alive today. George Knox examines two main questions: How can we analyse whether malformed babies cluster in time? If they do, what does it mean?

Clustering is usually defined with reference to time and/or spatial dimensions. One or the other of these dimensions may be implicit, as in this paper which looks at temporal clustering within cities as the implicit spatial unit of interest. It is useful to consider what is so special about time and space as risk factors that they should require their own statistical methodology. We can cope relatively easily, using classical statistical methods, with order in time and space, such as looking for yearly trends or seasonal fluctuation in the oesophageal atresia data. Clustering analyses, however, are looking for risk factors that are non-randomly and non-regularly distributed in space or time, such as infection or sources of environmental pollution. In addition, any division of continuously distributed time and space into subunits is purely arbitrary unless we know something about the underlying variation in risk factors.

To do a clustering analysis at all suggests that we don't know much about what the risk factors may be or their spatio-temporal distribution. We want the distribution of cases in time and space to give us clues about what the risk factors may be and how they are distributed. Clustering analyses are therefore by nature mainly hypothesis-generating (beyond the testing of very general hypotheses about the role of infection, for example), and have the statistical problems, particularly that of multiple testing and post-hoc testing, of all studies with poorly specified a priori hypotheses. The multiple testing in clustering analyses is slightly different from the 'usual' multiple testing associated with considering a wide range of risk factors and outcomes and their associations. In this paper, for example, we are concerned with only one outcome and one risk factor, time. The multiple testing problem relates to the many different ways in which we can subdivide and describe time (and the different statistical methods associated with different ways of summarising distributions in time). Here Knox looks at cases per month, cases per week, and intervals between cases of less than two weeks, recognising that these are arbitrary divisions. Knox also recognises the

post-hoc nature of the enterprise, since the time intervals to be analysed were suggested by the data.

Most analyses of clustering today would incorporate some comparison of the distribution of cases and non-cases, since the population (births in this case) is rarely distributed randomly in space or time. Knox did not incorporate this in this paper, although in an analysis purely considering time, such as this one, the clustering effect being looked for among cases is far in excess of the likely fluctuations in the number of births over time.

Statistical methods now available for clustering fall into two main groups: those based on counts of cases in small areas or time periods, and those based on the exact locations of cases and non-cases in space or time.¹ The methods based on areas or time periods (such as month in this analysis) impose arbitrary boundaries that can obscure clustering which crosses these boundaries. Clustering at a scale larger than the chosen areas can be detected by autocorrelation tests. Tests based on exact locations often use 'windows' which rove over time, space, or both, counting the number of cases occurring within each window, in a similar way to the method used in this paper where Knox counts the number of cases within 0-6 days of each index case. Since the same cases (and non-cases) may appear in a number of different windows, the counts are non-independent, as Knox pointed out, but statistical tests are now available to allow for this non-independence.^{1,2,3} Whichever of the two groups of methods are used, the areas, periods, or windows can be fixed temporo-spatial units, or population density dependent units. Time can be defined in terms of months or in terms of number of successive births. In space pre-specified administrative boundaries can be used, or for exact location data a distance based method,⁴ or population density based method.⁵

Knox also points to another problem, that of removing the yearly trend and potential seasonal fluctuation effects before looking for smaller scale clustering. This is part of a more general problem of removing the temporal and spatial variability in known risk factors before looking for clustering. Where tests for clustering are based on counts in predefined areas/periods, this can perhaps most easily be done by converting population numbers (eg, no births) to expected numbers of cases, and adjusting these expected numbers using classical

Department of Public Health and Policy, Environmental Epidemiology Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT
H Dolk

Correspondence to:
Dr H Dolk

Accepted for publication
October 1996

statistical methods for other covariates. Where tests are based on exact locations, controls can be chosen matched for the covariates of interest,⁵ or covariates can be entered directly into the statistical model.⁴

The availability of statistical tests for clustering has mushroomed since Knox's paper, including Knox's own contributions.^{6,7} Since research activity has been dominated by looking at cancers which have a long 'latent period' between exposure and diagnosis relative to congenital malformations, more emphasis has been placed on the spatial dimensions of clustering than on clustering in time. Moreover, the greater facility of handling large databases of spatially and temporally referenced data, and the greater availability of such data from large geographical areas means that analyses are more likely to concern space-time clustering than time clustering alone. Analyses in time are dominated by the needs of malformation monitoring, which seeks to detect increases in malformation frequency rapidly and with high sensitivity and specificity, and although the methodology overlaps with that of clustering analysis, prospective monitoring is statistically quite different.^{8,9}

The two principal groups of risk factors of interest in clustering research are infectious agents and environmental differences which tend to be localised in space or time (such as pollution of air or water). Little work has been done on defining and differentiating the clustering patterns that might be associated with either group of risk factors. Knox suggests a possible infectious origin for tracheo-oesophageal atresia in this paper, and has since postulated an infectious origin for hydrocephaly, anophthalmia, and other malformations which show evidence of clustering or epidemicity.¹⁰ The coincidental occurrence of epidemics of rubella and congenital rubella syndrome was one of the first instances where it can be said that investigation of time clustering led to identification of a risk factor for congenital malformations.¹¹ However, little progress has been made on the nature of the potential infectious basis for clustering of a range of malformations including tracheo-oesophageal atresia.

Environmental pollution tends to be a central concern in the investigation of clustering today, driven partly by heightened public concern. Recent concerns in relation to congenital malformations have included media reports of clusters of an/ microphthalmia and their potential link to pesticide exposure,¹² and a cluster of limb defects in the Isle of Wight which led to speculation about some 'coastal' risk factor.¹³ The evaluation of individual clusters is a different matter statistically to the evaluation of clustering as discussed above.¹ A post-hoc statistical probability attached to an individual cluster which gives the likelihood that it occurred by chance is at best only a guide as to whether it is worth spending resources investigating, but has no real statistical validity. In this sense we can question Knox's remark in his introduction that Gross' photograph of six affected babies in one hospital 'supplies a

concise statement of the specific problem with which this paper is concerned'. The paper is concerned with clustering, not the significance of any one particular cluster. Moreover, a rather unsatisfactory aspect of clustering analysis is that it does not necessarily allow the identification of individual 'true' clusters.

An increasingly important issue in the evaluation of clusters and clustering is the definition of the diagnostic group of interest. In this paper, Knox first establishes that it is reasonable to consider all cases of oesophageal atresia and tracheo-oesophageal fistula as one entity. Subsequently the tendency in clinical and epidemiological research into congenital malformations has been to demonstrate pathogenetic and aetiologic heterogeneity among cases of the 'same' malformation, a clear example being the 10% or so of oesophageal atresia cases which are associated with a chromosomal anomaly (not discovered in 1959). There is now much discussion¹⁴ as to whether for any particular malformation (like oesophageal atresia) children with multiple malformations should be considered separately from those with isolated malformations (ie splitting cases into subgroups), or whether apparently different malformations should be regrouped on embryologic or pathogenetic grounds (eg those due to vascular disruption, or those occurring in structures derived from cranial-neural-crest cells). A particularly intriguing cluster was reported in Hungary,¹⁵ apparently related to a chemical poisoning incident from fish ingestion, including cases of twins, Down syndrome, congenital heart disease, anal atresia, choanal atresia, cleft lip, and Robin sequence. In considering the issue of diagnostic groupings in relation to teratogens and teratogenic mechanisms, we can hope that clustering analyses and other epidemiological and clinical research will continue to be mutually informative.

- 1 Alexander FE, Cuzick J. Methods for the assessment of disease clusters. In: Elliott P, Cuzick J, English D, Stern R, eds. *Geographical and environmental epidemiology*. Oxford: OUP, 1992.
- 2 Wallenstein S. A test for detection of clustering over time. *Am J Epidemiol* 1980; **111**:367-71.
- 3 Weinstock M. A generalised scan statistic test for the detection of clusters. *Int J Epidemiol* 1981; **10**:289-93.
- 4 Diggle P, Chetwynd A. Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics* 1991; **47**:1155-63.
- 5 Cuzick J, Edwards R. Tests for spatial clustering in heterogeneous populations. *J Roy Stat Soc, Series B*; 1990; **52**: 73-104.
- 6 Knox EG. The detection of space-time interactions. *Applied Statistics* 1964; **13**:25-29.
- 7 Knox EG, Lancashire R. Detection of minimal epidemics. *Stat Med* 1982; **1**:183-89.
- 8 Chen R, McDowall M, Terzian E, Weatherall JAC. *EU-ROCAT guide to monitoring methods for malformation registers*. Brussels: Commission of the European Communities, 1983.
- 9 Gallus G, Mandelli C, Marchi M, Radaelli G. On surveillance methods for congenital malformations. *Stat Med* 1986; **5**:565-71.
- 10 Knox EG, Lancashire R. *The Epidemiology of congenital malformations*. London: HMSO, 1991.
- 11 Gregg NM. Congenital cataract following German measles in the mother. *Trans Ophthalmol Soc Aust* 1941; **3**:35-46.
- 12 Dolk H, Elliott P. Evidence for clusters of anophthalmia is thin. *BMJ* 1993; **307**:203.
- 13 Botting B. Limb reduction defects and coastal areas. *Lancet* 1994; **343**:1033-4.
- 14 Khoury M J, Botto L, Mastroiacovo P, Skjaerven R, Castilla E, Erickson J D. Monitoring for multiple congenital anomalies: an international perspective. *Epidemiol Rev* 1994; **16**:335-47.
- 15 Czeizel A, Elek C, Gundy S, et al. Environmental trichlorfon and cluster of congenital abnormalities. *Lancet* 1993; **341**: 539-42.