A B S T R A C T

Reports of disease clustering are becoming ever more common, and there is increasing pressure on public health agencies to respond rapidly and appropriately to these reports. We investigated a cluster of five cases of MS occurring in a small multidisciplinary hospital department. Data were collected by a variety of methods, including measurement and description of the workplace, completion of survey instruments by department staff, and construction of case histories of persons with MS. The results indicated that the department comprised a high-risk population and that only one case of MS could have any possible etiologic significance. Investigators should consider a number of factors when evaluating disease clusters, including the accuracy of diagnosis, the induction period and cause of the disease, and possible biases in the population at risk. Additionally, boundaries should not encircle the cases that led to identification of the cluster and should reflect environmental significance.

A B R É G É

Les rapports sur les agrégats de cas sont de plus en plus courants, et de plus en plus de pressions s'exercent sur les services de santé publique pour qu'ils répondent rapidement et comme il faut à ces rapports. Nous avons étudié un agrégat de cinq cas de sclérose en plaques survenus dans un petit service hospitalier, pluridisciplinaire. Les données ont été recueillies à l'aide de diverses méthodes, notamment la mesure et la description du lieu de travail, les réponses du personnel à des questionnaires et les dossiers d'observation médicale des personnes atteintes par la sclérose en plaques. Les résultats ont révélé que le service comprenait une population à risque élevé et qu'un seul cas pourrait avoir n'importe quelle explication causale. Les enquêteurs devraient prendre en considération un ensemble de facteurs lorsqu'ils évaluent des agrégats de cas, notamment l'exactitude du diagnostic, la période d'induction et la cause de la maladie, ainsi que les biais possibles dans la population à risque. En outre, il ne faut pas se circonscrire aux cas qui ont permis l'identification de l'agrégat mais prendre en compte l'ensemble de l'environnement.

Responding to Reported Clusters of Common Diseases: The Case of Multiple Sclerosis

Luanne M. Metz, MD, FRCPC, Sandra McGuinness, MN

One might suppose that the likelihood of two or more persons in a group having the same birthday would be extremely modest, yet, surprisingly, in a group of 23 people there is a 50% chance of just such an occurrence and, in a group of 47 people, the likelihood of concurrent birth dates rises to 95%.¹ The significance of this seemingly irrelevant piece of statistical trivia is that randomness does not imply uniformity.² Thus, although the occurrence of disease may be a random event, distribution of disease may not be consistent, and clusters of disease may arise by chance alone. Generally, however, among health professionals and the lay public alike, clusters of disease are perceived to have etiologic significance following the simple train of logic that a high concentration of disease must imply a similar concentration of the causative agent.³⁻⁶ In recent years public health agencies have been besieged by reports of clusters of disease in communities, workplaces and schools.⁷⁻⁹ Most commonly, these reports are accompanied by requests for scientific studies both to unlock the etiologic significance of the cluster and to protect the, as yet, unexposed public from the development of future disease. We present a case study of a cluster of multiple sclerosis (MS) that occurred in a large Canadian hospital and suggest methods for managing cluster reports and requests for scientific studies gleaned from our own experience and that of other researchers before us.

Case study: A cluster of MS

A cluster of five cases of MS occurring in a small, multidisciplinary hospital department was reported to the director of a university affiliated neurology clinic in Western Canada by a staff member who had recently been given a diagnosis of MS. In response to administrative and staff requests, we conducted an exploratory descriptive study of the staff and the workplace.

Financial constraints limited the extent of the study to the hospital department originally reporting the MS cluster. Although self-contained, this department was close to other hospital departments and shared common air and water supplies with several inpatient units. No hazards unique to the department were identified. All current and previous employees (to the last eight years) were invited to participate in the study. Data were collected by a variety of methods, including measurement and description of the workplace environment (air and water quality and cobalt levels), completion of survey instruments by department staff, and construction of case histories of persons with MS. Hazard evaluation was conducted over an eight month period to allow repeated measures over the summer and winter seasons. Concurrently, a literature review was completed with special attention paid to predisposing factors associated with MS, the incidence, prevalence and epidemiology of MS, and suspected causal factors.

Detailed data were available on 15 of the 23 eligible staff members. All 15 participants were female Caucasians from high-risk ancestry groups who were resident in areas of high environmental risk for MS during the putative period of acquisition of MS risk (that is, prior to age 20).¹⁰⁻¹⁶ Only

University of Calgary Multiple Sclerosis Clinic, Calgary General Hospital

Correspondence and reprint requests: S. McGuinness, MS Clinic Room AC137A, SSB, Foothills Hospital, 1403 29 Street NW, Calgary, AB, T2N 2T9, Tel: 403-670-4246, Fax: 403-283-2270, E-mail: sdmcguin@acs.ucalgary.ca

persons with MS had a family history of disease. No common environmental factors operant in the department that could account for the occurrence of MS were found, and there were no significant differences in employment, health history, or workplace exposures between persons with MS and without MS.

Of the five persons identified as having MS, three had clinically definite MS, while for two the diagnosis was possible MS only.¹⁷ Two of the three occurrences of definite MS were found to be etiologically non-significant; that is, in one case the onset of MS was one month after the start of employment and in the other it was 18 years prior to employment.^{11,15,16} There remained one occurrence of definite MS that could have possible etiologic significance. In this admittedly high-risk population, we were unable to conclude that the occurrence of MS in one individual was indicative of hazardous workplace exposures. In the year after the study, no further occurrences of either possible or definite MS have appeared, and MS has been ruled out in one of the two individuals with possible MS.

Principles of cluster investigation

The US Centers for Disease Control defines a cluster as "an unusual aggregation, real or perceived, of health events that are grouped together in time and space and that are reported to a health agency."¹⁸ This definition integrates several important concepts. First, although a distinction is implied between perceived and real clusters, the relevance of both real and perceived clusters is recognized. As in our experience, when fully examined many clusters reported to health agencies are found not to be clusters at all.^{19,20} This commonly occurs in reported cancer clusters, when more than one type of cancer is grouped into the reported cluster.¹⁻² However, regardless of the true significance of the cluster, the aggregation of health events is important to the involved individuals, and some official action by the health agency is required.⁸ Second, this definition raises the distinction between temporal and spatial clustering. When evaluating clusters to determine causal significance the likely or known temporal and causal

sequence of disease should be considered. In diseases with long induction periods, such as MS, temporal and spatial clustering will have causal significance only when such clustering occurs during the induction period. Temporal and/or spatial clustering once the disease is evident does not have causal significance, and may be the result of many diverse factors.²¹ Last, by implication, this definition suggests that official responses to reported clusters are crucial.

When assessing the etiologic significance of any cluster, a number of factors should be considered:

1. Diagnosis must be accurate. Only those individuals who conclusively exhibit the disease in question should be counted as cases.²² In our example, removal of two individuals in whom MS was not confirmed reduced by 40% the size of the cluster.
2. The detection of temporal and/or spatial clustering may or may not have causal significance. As the statistician can easily illustrate, random distribution of events does not imply uniform distribution. Clusters of any event can and do occur by chance alone, particularly in the more common diseases, such as MS.^{1,2} Although the five MS cases reported to our department were clustered both temporally and spatially, this clustering was present only after onset of disease and had no etiologic significance.
3. Clusters of disease identified post hoc almost always include in the cluster those cases that originally led to identification of the cluster, thus a higher than expected rate of disease is predetermined.³⁻⁵ Essentially, one discovers that which has already been reported. Our investigation uncovered no new cases of MS beyond those already reported, and no new cases have appeared in the 18 months since completion of the study. Given the long induction period for MS, surveillance of our hospital department would need to continue for an additional 20 years to accurately establish the rate of disease. Few health departments have budgets to support this type of long-term surveillance.
4. The foregoing point illustrates the importance of careful consideration of the boundaries that will be drawn around the population. Boundaries that encircle only the cases that led to the identification of the cluster are guaranteed to demonstrate higher than expected prevalence and may not take account of induction periods or biological plausibility.²² The recently completed landmark cross-Canada genetic study of MS indicated that the environmental factors influencing susceptibility to MS are not operant at a micro-environmental level; rather they are influential across broad regional populations.²³ We found no evidence to indicate that the environment to which staff in the department had been exposed differed significantly from that of any other department in the hospital, or indeed in the local community. From an environmental exposure point of view, the boundary should have encircled, at a minimum, the entire hospital and neighbouring community. The population, then, would encompass not only members of the department in question, but all hospital employees, past and present, and patients and visitors. Unfortunately, such a large investigation was beyond our budget, and is likely beyond the budget of most other comparable health agencies.
5. Clusters of under five cases are not large enough to support traditional epidemiologic statistical analysis.^{8,22} Although it may be tempting to compare the prevalence in the cluster population with the expected prevalence, such comparisons are not legitimate, since the true denominator in the cluster population may be difficult or impossible to determine. In our example, a ratio of 3 cases to 23 staff results in a prevalence of 13,043 per 100,000, clearly much in excess of that expected in the region. However, it is unlikely that 23 represents the true denominator of all those potentially exposed. The true denominator, in fact, cannot be determined. It may be the entire staff of the hospital, past and present, as well as all potentially

exposed patients and visitors. Additionally, the confidence intervals around the prevalence rates of small clusters in small populations will be large, and thus the prevalence rate itself is imprecise.^{24,25}

6. The experience of previous cluster studies suggests that the investigation of reported clusters is unlikely to uncover factors of etiologic significance unless a rare disease is clustered in a population without positive bias and in which a single causal agent is implicated with a dose-response relationship to the disease in question. None of these conditions was met in our MS cluster. Susceptibility to MS is the result of interaction of a number of factors, including broad environmental exposures and genetic predisposition. In our region, the prevalence rate for MS is in the vicinity of 216 per 100,000, and females outnumber males by a factor of two.²⁶ There are many reports of MS clustering in the literature, and although some have been explained by socio-economic factors, the cause of the majority remains obscure.²⁷⁻³¹ Given the post-hoc nature of cluster recognition, it is likely that most unexplained clusters are simply the result of chance occurrences of disease without etiologic significance.

The foregoing is not meant to suggest that the investigation of clusters of commonly occurring diseases is an entirely fruitless enterprise. On the contrary, it is our experience that sensitive, appropriate and timely official responses to cluster reports are highly desirable, and do much to discharge anxiety and distress. Rather, we suggest these points as some factors that should be considered in calls for investigation of any reported cluster of disease. In our MS cluster, thorough exploration of the reported cluster was sufficient to ameliorate staff anxiety about workplace conditions, because we were able to demonstrate that the occurrence of MS was without etiologic significance.

CONCLUSIONS

Fortuitous identification of disease clusters has granted epidemiologists and public health personnel considerable success in

differentiating disease-causing agents in communities and workplaces. Yet not all reported clusters have etiologic significance. The role of random chance in the distribution of disease should not be overlooked nor should the morass of other factors (such as genetic predisposition and socioeconomic elements) that are influential in determining the prevalence of disease. As individuals and communities become more proactive in health-related matters it is likely that public health agencies will receive an ever increasing number of cluster reports and the accompanying requests for scientific studies. Disease clusters are generally associated with a great deal of anxiety and distress among those who may or may not be affected; therefore, sensitive, appropriate and timely official responses are required and do much to discharge anxiety and distress. However, a number of factors should be considered when investigating reported disease clustering, and the etiologic significance of any reported cluster should not be overestimated. Conversely, the investigation of the reported cluster may lead to the identification of an important causative agent, and thus the opportunity for improving the fate of human kind is always tantalizingly close.

REFERENCES

1. Taylor D, Wilkie D. Drawing the line with leukaemia. *New Scientist* 1988; 21 July 53-6.
2. Enterline PE. Evaluating cancer clusters. *J Am Ind Hyg Ass* 1985;46:B10-B13.
3. Sheremeta WA, Poskanzer DC, Withum DG, MacLeod CL, Whiteside ME. Unusual occurrence on a tropical island of multiple sclerosis. *Lancet* 1985;8455: 618.
4. Dean G, McDougall EI, Elian M. Multiple sclerosis in research workers studying swayback in lambs: An updated report. *J Neuro, Neurosurg Psychiatry* 1985;48: 859-65.
5. Stein EC, Schiffer RB, Hall WJ, Young N. Multiple sclerosis and the workplace: Report of an industry-based cluster. *Neurology* 1987;37:1672-77.
6. Ingalls TH. Clustering of multiple sclerosis in Galion, Ohio, 1982-1985. *Am J Forensic Med Pathol* 1989;10:213-15.
7. Schulte PA, Ehrenberg RL, Singal M. Investigation of occupational cancer clusters: Theory and practice. *Amer J Pub Health* 1987;77:53-5.
8. Fiore BJ, Hanrahan LP, Anderson HA. Public health response to reports of clusters. *Am J Epidemiol* 1990;132: S14-S22.
9. Bender AP, Williams AN, Johnson RA, Jagger HG. Appropriate public health responses to clusters: The art of being responsibly responsive. *Am J Epidemiol* 1990;132:S48-52.
10. Compston A. The epidemiology of multiple sclerosis: Principles, achievements and recommendations. *Ann Neurol* 1994;36:S211-S217.
11. Kurtzke JF. Epidemiologic evidence for multiple sclerosis as an infection. *Clin Microbiol Rev* 1993;6:382-427.
12. Page WF, Kurtzke JF, Murphy FM, Norman JE. Epidemiology of multiple sclerosis in US veterans: V. Ancestry and the risk of multiple sclerosis. *Ann Neurol* 1993;33:632-39.
13. Sadovnick AD, Ebers GC. Epidemiology of multiple sclerosis: A critical overview. *Can J Neurol Sci* 1993;20:17-29.
14. Weinshenker BG, et al. The natural history of multiple sclerosis: A geographically based study. *Brain* 1989;112:133-46.
15. Kurtzke JF, Hyllested K. Multiple sclerosis in the Faroe Islands: 1. Clinical and epidemiological features. *Ann Neurol* 1979;5: 6-21.
16. Poskanzer DC, Walker AM, Prenney LB, Sheridan JL. The etiology of multiple sclerosis: Temporal-spatial clustering indicating two environmental exposures before onset. *Neurology* 1982;31:708-13.
17. Poser CM, et al. New diagnostic criteria for multiple sclerosis: Guidelines for research protocols. *Ann Neurol* 1983;13:227-31.
18. Centers for Disease Control. Guidelines for investigating clusters of health events. *MMWR* 1990;39:RR1-RR11.
19. Hoffman RE, Zack MM, Davis LE, Burchfiel CM. Increased incidence and prevalence of multiple sclerosis in Los Alamos County, New Mexico. *Neurology* 1981;31:1489-92.
20. Conomy JP, et al. Multiple sclerosis in Galion, Ohio: A cluster in time or a cluster in space? *Neurology* 1989;39:284.
21. Kreiger N, Reynolds D. An investigation of two apparent cancer clusters in one community. *Can J Public Health* 1992;83:116-19.
22. Rothman KJ. A sobering start for the cluster busters' conference. *Am J Epidemiol* 1990;132:S6-S13.
23. Ebers GC, et al. A genetic basis for familial aggregation in multiple sclerosis. *Nature* 1995;377:103-4.
24. Hibberd PL. Use and misuse of statistics for epidemiologic studies of multiple sclerosis. *Ann Neurol* 1994;36:S218-S230.
25. Klein GM, Rose MS, Seland TP. A prevalence study of multiple sclerosis in the Crownsnest Pass region of Southern Alberta. *Can J Neurol Sci* 1994;21:262-65.
26. Svenson LW, Woodhead SE, Platt GH. Regional variations in the prevalence rates of multiple sclerosis in the Province of Alberta, Canada. *Neuroepidemiology* 1994;13:8-13.
27. Helmick CG, et al. Multiple sclerosis in Key West, Florida. *Am J Epidemiol* 1989;130:935-38.
28. Hader WJ, Irvine DG, Schiefer B. A cluster-focus of multiple sclerosis at Henribourg, Saskatchewan. *Can J Neurol Sci* 1990;17:391-94.
29. Eastman R, Sheridan JL, Poskanzer DC. Multiple sclerosis clustering in a small Massachusetts community with possible common exposure 23 years before onset. *Med Intell* 1973;289:793-94.
30. Stein EC, Schiffer RB, Hall WJ, Young N. Multiple sclerosis and the workplace: Report of an industry-based cluster. *Neurology* 1987;37:1672-77.
31. Murray TJ. An unusual occurrence of multiple sclerosis in a small rural community. *Can J Neurol Sci* 1976;3:163-66.

Received: October 17, 1996

Accepted: February 10, 1997