Background: The spatial and temporal pattern of excessive disease occurrence, termed clustering, may provide clues about the underlying etiology.

Objective: To report the occurrence of 3 clusters of Parkinson disease (PD) in Canada.

Design and Patients: We determined the population groups containing the clusters, geographical limits, and duration of exposure to the specific environments. We tested whether there was an excessive presence of Parkinson disease by calculating the probability of the observed cases occurring under the null hypothesis that the disease developed independently and at random in cluster subjects. Results of genetic testing for mutations in the α-synuclein, parkin, tau genes, and spinocerebellar ataxia genes (SCA2 and SCA3) were negative.

Results: The probabilities of random occurrence (P values) in the 3 clusters were P = 7.9 × 10⁻⁷ for cluster 1, P = 2.6 × 10⁻⁷ for cluster 2, and P = 1.5 × 10⁻⁷ for cluster 3.

Conclusions: Our findings indicate an important role for environmental causation in Parkinson disease. A possible role exists for environmental factors such as viral infection and toxins in the light of current evidence.

Arch Neurol. 2004;61:1057-1060

The grouping of cases of a particular disorder in space and time, termed clustering, may provide clues about the underlying etiology. Only a few reports on clustering of Parkinson disease (PD) have been published.1,2 We herein report 3 PD clusters in Canada and review current evidence for PD causation in the light of our findings.

METHODS

DESCRIPTION OF THE CLUSTERS

All subjects, except 1 inaccessible individual in cluster 1, underwent examination by us (D.B.C., A.K., and S.M.C.). We confirmed the presence of PD based on clinical diagnostic criteria.3 Parkinson disease in all subjects was typical with respect to clinical features and levodopa responsiveness. Detailed family and occupational histories were obtained. An epidemiologist (C.V.N.) examined the working environments. The subjects, with the exception of the inaccessible individual in cluster 1, underwent genetic testing for known parkin and α-synuclein gene mutations, and mutations in the tau (exons 1, 7, and 9-13), and spinocerebellar ataxia genes (SCA2 and SCA3). The 3 subjects who were older than 50 years at onset of symptoms did not undergo testing for the parkin mutation.4 The only subject with a family history of PD had negative test results for mutations in the α-synuclein gene.

Cluster 1

This cluster consisted of 4 subjects, including 3 men and 1 woman aged 42 to 65 years (mean ± SD age, 57.0 ± 9.4 years), who had worked together as part of a television crew of approximately 125 people (male-female ratio, 5:2) for 5 years 30 years ago. The age at PD onset ranged from 30 to 56 years (mean ± SD age, 44.8 ± 12.8 years). Symptoms developed in all patients within 11 years of each other. None of the cases had a family history of PD. One subject had a bout of pneumonia that was probably viral 10 years before symptom onset. The production crew worked in poorly ventilated underground studios. Testing of air samples from the studios revealed high concentrations of carbon dioxide. The period of exposure to this environment was 5 years in all subjects. There was a latent period of 3 to 13 years (mean ± SD, 9.5 ± 4.5 years) from cessation of exposure to this environment to the development of symptoms.
Cluster 2

This cluster consisted of 4 subjects, all men aged 59 to 75 years (mean ± SD, 65.0 ± 7.0 years), who had taught at the same college for periods ranging from 19 to 30 years. Their ages at PD onset ranged from 52 to 70 years (mean ± SD years, 58.5 ± 8.1 years). Symptoms developed in each subject within 8 years. None had a family history of PD. Two subjects had been intermittently exposed to organic solvents such as benzene and toluene. Another subject had experienced an episode of encephalitis in childhood. All subjects had negative results on genetic testing. The subjects had worked for periods ranging from 7 to 12 years (mean ± SD, 9.8 ± 2.1 years) in a portable classroom situated over a filled-in waste dump. There were never more than 30 other people (male-female ratio, 5:1) in the trailer on a regular basis, inclusive of the 4 subjects. The latent period between cessation of exposure and development of symptoms ranged from 12 to 18 years (mean ± SD, 15.8 ± 2.6 years).

Cluster 3

This cluster consisted of 3 subjects, including 2 men and 1 woman aged 51 to 75 years (mean ± SD age, 60.0 ± 13.1 years), who were among 7 employees (male-female ratio, 4:3) working together in the office of a small garment-manufacturing factory. The age at PD onset ranged from 30 to 71 years (mean ± SD age, 47.0 ± 21.4 years). The period of symptom onset from the first to the last patient was 17 years. One subject’s father had PD. One subject had an episode of high fever and confusion of uncertain cause about 6 years before symptom onset. No history of toxic exposure was found. Results of genetic testing were all negative. The subjects had worked together in the same office for periods ranging from 2 to 35 years (mean ± SD, 15.7 ± 17.2 years) and developed PD while working in this environment. Details of the 3 clusters are summarized in the Table.

STATISTICAL ANALYSIS

It is difficult to demonstrate directly that a group of individuals constitute a cluster in time and space with respect to a disease. In addition to the obvious problem of ascertainment bias associated with the retrospective identification of such groups and the difficulty of defining a suitable control framework for the purposes of comparison, there is a major logical barrier to our attempt to prove the hypothesis that the group constitutes a meaningful cluster. As Popper stated, “Every scientific hypothesis must be testable, and the way to test it is to look for circumstances in which it does not hold.”23 It is therefore only by attempting to disprove the null hypothesis that a disease developed independently and at random in a group of individuals that the likelihood of a true cluster may be scientifically tested.

We postulated the null hypothesis that for each of the 3 groups examined, the development of PD occurred independently among subjects and followed the age- and sex-specific population incidence rates of PD appropriate to their locations and times. The age- and sex-specific incidence rates were derived from the Rochester, Minn, study.6 These rates reflect the occurrence of PD in other parts of the North American continent and are comparable to those of other studies.27 Because some of our PD cases were young at PD onset, we smoothed the published incidence rates at the early ages by means of cubic splines, applying a moderate tension of 3.28 The spline-based approximations reproduced the reported incidence rates extremely well but provided smoother and more precise estimates for the early ages.

Our calculations proceeded in analogy with the classic statistical testing method of the null hypothesis that a coin toss is unbiased when we suspect that it may in fact have a bias toward coming up heads. We test the coin by observing the number of occurrences of heads in n tosses and by calculating the P value (ie, the probability of obtaining as many heads as observed, or even more, under the null hypothesis that the coin toss is unbiased so that the observed heads occurred independently and at random). Using the age-specific, spline-based, and sex-specific incidence rates, we calculated for each group the probability (P value) under the null hypothesis that as many (or more) PD cases as were observed in that particular group could have happened independently and at random at any time during the entire follow-up period of each individual (from onset of exposure to present) in that group. In the case of the subject with a positive family history in cluster 3, we took into account the 4 × higher-than-normal risk for disease occurrence29 and made the appropriate correction.

In cluster 1, PD developed in 4 individuals (3 men and 1 woman), with a mean exposure-to-onset time of 14.5 years. The remaining 121 subjects (male-female ratio, 5:2) were followed up for a mean of 27.0 disease-free years. Under the null hypothesis, the probability of development of PD in 4 or more individuals from this group during the observed follow-up was $P = 7.9 \times 10^{-7}$. In the same cluster, if PD developed in 3 or more individuals, the corresponding null hypothesis probability would have been $P = 4.8 \times 10^{-3}$; if it developed in 2 or more, $P = 2.2 \times 10^{-1}$; and if it developed in 1 or more, $P = .07$, thus reaching nonsignificance at the conventional level of $P < .05$.

In cluster 2, PD developed in 4 men, with a mean exposure-to-onset time of 25.5 years. The remaining 26 subjects (male-female ratio, 5:1) were followed up for a mean of 31.8 disease-free years. Under the null hypothesis, the probability of development of PD in 4 or more individuals from this group during the recorded follow-up intervals was $P = 2.6 \times 10^{-7}$.

In cluster 3, consisting of 7 subjects (4 men and 3 women), PD developed in 3 (2 men and 1 woman), with a mean exposure-to-onset time of 15.7 years. The remaining 4 subjects were followed up for a mean of 21.5 disease-free years. Under the null hypothesis of independent events occurring at random, the probability of development of the disease in 3 or more individuals during the observed follow-up was $P = 1.5 \times 10^{-7}$.

### Table

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of incident cases</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Sex ratio, No. M/F</td>
<td>3.1</td>
<td>4</td>
<td>2:1</td>
</tr>
<tr>
<td>Age, y</td>
<td>57.0 ± 9.4</td>
<td>65.0 ± 7.0</td>
<td>60.0 ± 13.9</td>
</tr>
<tr>
<td>Age at exposure, y</td>
<td>30.2 ± 10.8</td>
<td>33 ± 8.7</td>
<td>31.3 ± 9.9</td>
</tr>
<tr>
<td>Exposure, y</td>
<td>5.0 ± 0</td>
<td>9.75 ± 2.1</td>
<td>15.7 ± 17.2</td>
</tr>
<tr>
<td>Age at symptom onset, y</td>
<td>44.7 ± 12.8</td>
<td>58.5 ± 8.1</td>
<td>47 ± 21.4</td>
</tr>
<tr>
<td>Latent period, y</td>
<td>9.5 ± 4.5</td>
<td>15.7 ± 2.6</td>
<td>NA</td>
</tr>
<tr>
<td>No. of subpopulation (sex ratio, No. M/F)</td>
<td>125 (5:2)</td>
<td>30 (5:1)</td>
<td>7 (4:3)</td>
</tr>
<tr>
<td>Family history</td>
<td>None</td>
<td>None</td>
<td>PD in 1 subject’s father</td>
</tr>
</tbody>
</table>

Abbreviations: NA, not applicable; PD, Parkinson disease.

*Data are given as mean ± SD unless otherwise indicated.
There has been controversy about the relative importance of genetic vs environmental factors in causing PD. Genetic and environmental factors are not mutually exclusive, as genetic susceptibility may confer selective vulnerability to risk factors. Although identified mutations account for a minority of patients with PD, in most patients the cause remains unknown. The cluster approach, although useful in determining the underlying etiology in the case of acute illnesses, has been less rewarding for chronic illnesses, especially when there is no plausible biological explanation for grouping of cases. Despite some skepticism, we consider that clusters may yield epidemiologically useful information that we can ill afford to ignore.

The odds of our 3 clusters of PD occurring by chance alone are very low. The absence of mutations known to be associated with PD together with the occurrence of the young-onset form in 4 patients lends support to environmental causation. A large survey comparing monozygotic and dizygotic twins failed to reveal a higher concordance rate in the monozygotic group in the age range when PD usually starts, supporting environmental causation. Furthermore, familial occurrence does not necessarily mean genetic causation, because family members share their environment and their genes. One study demonstrated that the risk for development of PD in a child from a parent-child cluster depended on the child’s age when the parent started to show symptoms rather than the parent’s age; younger children had greater risks. This finding suggests environmental risk factors. A prominent feature was the long latent period between cessation of exposure to the shared environment and onset of PD symptoms. This pattern suggests an environmental cause that exerts its influence during a short time during the period of the shared environment and results in a cascade of events culminating in the clinical manifestations of PD after a long delay. One study used a mathematical model based on observations relating to clinical deficits to indicate the most likely pathogenesis is an event that kills some neurons and damages others in such a way that their life expectancy is reduced, or an event that causes a mechanism that kills healthy neurons at a constant rate. These models best explain the occurrence of PD clusters such as ours. The long latent period is in keeping with the well-recognized fact that at least 50% of nigral neurons must be lost before the symptoms of PD become clinically manifest.

What then is the character of the event or the events that result in the death of nigral neurons? Toxic and infective causes are prime candidates. The discovery that 1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine (MPTP) leads to selective destruction of the nigrostriatal pathway opened up a new vista of thought on the potential role of toxins. However, attempts to link MPTP to ordinary PD have been unsuccessful. Several other controversial reports have described the potential role of herbicides, fungicides, and pesticides. Speculation about the role of infection dates back to the epidemic of von Economo encephalitis in which parkinsonism was a sequel, sometimes after a delay of several years. Progression of parkinsonism has been documented in these patients, despite absence of markers of persisting viral infection.

Positron emission tomographic scanning of the brain has demonstrated selective lesions of the nigrostriatal pathway after viral encephalitis. The pattern of striatal dopamine terminal loss is highly reminiscent of PD, with maximal involvement of the posterior putamen. Japanese workers have shown that certain strains of influenza virus A are selectively tropic to the nigral neurons and that some gain access to the brain via the nasal passages in mice. Another study showed that antibodies to the Epstein-Barr virus cross react with α-synuclein in the brains of patients with PD. Although no evidence of ongoing viral infection in PD has been reported, atypical inflammatory reactions have. The viral hypothesis is also buttressed by an epidemiological report that the prevalence of PD is more than twice normal in teachers, medical workers, loggers, and miners. The simplest explanation for this finding is an infectious origin. The increased risk for teachers and medical workers is obvious. Most loggers and miners, in the relevant time frame of this study, shared cramped sleeping quarters that put them at increased risk for respiratory infections. Another study examined occupational risk factors in monozygotic twins discordant for PD, and also found a significantly increased risk associated in teachers and health care workers. This finding is particularly cogent because genetic confounding factors are eliminated as contributing risk factors.

A tenable explanation for the occurrence of PD clusters would be a brief infective or toxic exposure (by no means mutually exclusive) in the shared environment, resulting in PD several years later. Advances in virology and toxicology, together with old-fashioned epidemiological sleuthing, hold the key to unraveling the origins of PD.

Accepted for publication February 5, 2004.

Author contributions: Study concept and design (Drs Kumar, Schulzer, Van Netten, Tsui, and Calne); acquisition of data (Drs Kumar, Van Netten, and Calne); analysis and interpretation of data (Drs Kumar, Schulzer, Wszolek, Van Netten, Tsui, Stoesell, and Calne); drafting of the manuscript (Drs Kumar, Schulzer, and Tsui and Ms Calne); critical revision of the manuscript for important intellectual content (Drs Kumar, Schulzer, Wszolek, Van Netten, Stoesell, and Calne); statistical expertise (Dr Schulzer and Mr Mak); obtained funding (Drs Stoesell and Calne); administrative, technical, and material support (Ms Calne and Drs Wszolek, Van Netten, Stoesell, and Calne); study supervision (Drs Kumar, Wszolek, Van Netten, Tsui, Stoesell, and Calne).

This study was supported by the Canadian Institutes for Health Research (Dr Stoesell); Canada Research Chairs, Ottawa, Ontario; National Parkinson Foundation, Miami, Fla (Ms Calne); Pacific Parkinson’s Research Institute, Vancouver (Dr Calne); and Parkinson Society British Columbia, Vancouver (Dr Schulzer).

We thank Matt Farrer, PhD, and M. Hutton, PhD, for the genetic testing, and Sharon Yardley, RN, for her assistance with blood sampling.

(Reprinted) Arch Neurol/Vol. 61, July 2004

©2004 American Medical Association. All rights reserved.
REFERENCES


