When is a cluster of disease really a cluster?

Public health agencies often investigate what seems to be an excess of disease in a particular time and place, i.e. a disease cluster. The purpose is to control health hazards if they are causing disease in affected populations. Investigations are usually precipitated by public concerns, or increased risks detected by disease surveillance systems. This issue contains a report on the approach to occupational disease clusters taken by the UK Health & Safety Executive (HSE) [1]. In keeping with other health protection agencies [2,3], the HSE’s approach pragmatically combines systematic but flexible methods of population health measurement, judgement under uncertainty and precaution in ensuring that potential disease risks are not neglected.

Disease cluster investigations, however, tend usually to find nothing of interest [1–3]. Common reasons for this include small numbers of cases, imprecise case definitions, difficulties in defining populations at risk, vague exposure data and biased information resulting from publicity [3,4]. The HSE paper in this issue reports that ‘most workplace clusters turn out to have chance as the most likely explanation’ [1]. Similarly, United States Centers for Disease Control (CDC) guidelines state that ‘some clusters occur by chance’ [2]. This seems paradoxical because, from a statistical perspective, if the frequency of disease is significantly increased, then by definition it is very unlikely to be due to chance. Central to these problems are ambiguities in the meaning of ‘chance’ and of ‘cluster’.

The paradox of chance clusters can partly be unravelled by bearing in mind what chance means. One interpretation is that, if several cases of disease occur together due to chance, this means that they are causally unrelated [4]. Causal inference requires judgement that is based on more than just statistical hypothesis testing. It also entails considering the strength of association, the presence of dose–response relationships, prior knowledge and so on. Furthermore, if clusters are defined to include subjective opinions, then it is unsurprising that causal relationships are often not found. For example, CDC’s definition of a cluster is ‘an unusual aggregation, real or perceived, of health events that are grouped together in time and space’ [2]. Clearly, perceptions are fallible and may be uncorroborated by more objective data. The HSE definition is more ambiguous: ‘a marked excess compared to that expected’ [1]. The HSE definition raises the question as to whether expectation is used in the same way as by statisticians, who objectively compare observed and expected health events, or whether it refers to the subjective expectations of the public, health agencies, the media and others. It probably invokes both meanings, but one needs to keep their distinction in mind.

There are many sophisticated methods for investigating disease clusters [2,3,5], but they are all founded on basic statistical and epidemiological principles [4,5]. A key problem is the possibility of two types of error. Type 1 error occurs when there appears to be an excess of disease, or to be an association between exposure and disease, when in fact no causal relationship exists. Type 1 error is more likely if multiple statistical tests are performed. By definition, one would expect to find $P < 0.05$ in 1 in 20 significance tests, even if no causal relationship exists. If one performs 10 tests, then there is a 40% [that is, $1 - (0.95^{10})$] chance that at least one would be significant at the 5% level. This is clearly a problem for disease surveillance systems that continuously look and test for excesses of disease. Most statistical methods for investigating clusters aim to reduce the risk of type 1 error in one way or another [2,3,5]. But from the point of view of risk-averse populations, the risk of missing a real hazard may be more important than the risk of detecting a non-existent one. Type 2 error occurs when no statistical association is found, but a causal relationship exists. Type 2 error may result from flawed disease or exposure data, or inappropriate study design. The commonest reason, however, is insufficient statistical power—the cases are just too rare. It may be impossible, for example, to know whether three cases of a rare disease in a population of workers really is excessive. General solutions include expanding the investigation to larger populations and more precisely characterizing disease cases, exposures and populations at risk. But one cannot entirely avoid the risk of type 1 and type 2 errors.

One solution is to avoid statistical significance tests and to focus instead on estimating the strength of association between exposure and outcome, i.e. the magnitude of the risk increase. Thus, one should aim to estimate relative risks, or similar indicators, with confidence intervals showing the precision of one’s estimates. From a public health perspective, it is at least as important to know whether the relative risk or risk difference is large or small as it is to know the $P$-value [4]. ‘Data dredging’ procedures, such as in routine surveillance of multiple diseases, are inherently hypothesis generating rather than hypothesis testing and so it is appropriate that estimation gets higher priority.

Prudence and subjective judgement on the part of health agencies are unavoidable. Bayesian statistical methods, which are mainly used in cluster investigations...
to avoid type 1 errors [2,3], can also be used to combine information on people’s subjective probabilities of risk with epidemiological data [6]. One can describe the prior beliefs about a causal relationship among sceptics and among ‘believers’. These prior probability distributions can be combined with epidemiological data to assess whether the evidence of a causal relationship is strong enough to convince a sceptic that a problem exists, or to reassure a believer. Epidemiologists and statisticians, however, differ among themselves as to whether subjective probabilities provide admissible evidence and the data may never convince a sceptic or a believer. But this still underdeveloped application of Bayesian methods to disease cluster investigation may be useful to health agencies who need to decide when to act.

Disease cluster investigation exemplifies public health practice as ‘the science and art of improving a population’s health through the organized efforts of society’ [7]. Pragmatism must, however, be based on clear scientific reasoning and on understanding the limitations and potential of continually developing epidemiological and statistical methods.

Max O. Bachmann
School of Medicine, Health Policy and Practice, University of East Anglia, Norwich NR4 7TJ, UK

References