Identifying the exposure or agent responsible for causing a particular disease, is often a prerequisite to developing effective prevention strategies, and improving patient management. However controversy may accompany the detection or refutation of a causal link. A good example was the debate in South Africa about the cause of acquired immunodeficiency syndrome (AIDS).

How should we approach causality? Robert Koch, famed for his discovery of Mycobacterium tuberculosis and Vibrio cholerae, two ancient organisms that remain a plague in the 21st century, proposed four criteria for establishing the causative agent of an infectious disease. These postulates may be summarised as:

- The organism should be isolated in pure culture from each case of the disease;
- The organism should not occur in any other disease as a chance and non-pathogenic occurrence;
- Once isolated it should be grown in a series of cultures;
- This cultured organism should reproduce the disease on inoculation into an experimental animal.

Although these criteria are useful, particularly for diseases with a bacterial origin, they have proved less valuable for most viral diseases. Although the human immunodeficiency virus (HIV) was isolated from an AIDS patient, soon after the description of the syndrome in five Californian males, there are a number of constraints to applying the experimental method for proving causality. These include the long incubation period between exposure and disease onset, unethical nature of a randomised double-blind trial in which uninfected human volunteers are exposed to HIV or placebo, and lack of a suitable animal model. An experiment may not be legitimate as Sir Austin Bradford Hill defined useful criteria for assessing whether a demonstrated association between a putative exposure and a disease can be judged as causal or not. We will briefly introduce these criteria in this article, and refer to the association between HIV and AIDS to illustrate each. Readers are challenged to explore this association further.

Bradford-Hill's first criterion is the strength of association. When considering the strength of association it is useful not only to consider the strength of the measure (relative risk or odds ratio) but the type of study done. This hierarchy of studies will be considered in more detail in a future article in this series. In our present example, sophisticated detection methods inevitably demonstrate viral genetic material, antigens, and the virus itself, in patients with the clinical AIDS syndrome. A case-control study that provided strong evidence of the strength of association, compared HIV-negative and HIV-positive blood recipients who had been given transfusions for similar diseases. At follow up, 37 cases of AIDS developed in the HIV-infected group, but not a single AIDS-defining illness was found in the HIV-seronegative transfusion recipients.

The second factor considered when judging an association for causality is the consistency of the association. When different researchers, in different geographical settings and at different times, demonstrate the same association, then causality becomes more likely. In every country where AIDS has occurred, researchers working in different laboratories and using a variety of testing techniques have demonstrated the presence of HIV.

The third criterion is self-explanatory - the temporal relationship of the association. The onset of clinical disease should follow after the exposure. To the authors' knowledge, no scientific report exists of an AIDS patient who was initially HIV negative but after a number of months or years of follow-up became positive. In developed countries the median period of time between infection with HIV and the onset of clinically apparent disease has been demonstrated to be approximately 10 years, in prospective studies of homosexual men in which dates of sero-conversion are known, HIV-infected blood-transfusion recipients, injection-drug users, and adult haemophiliacs.

An important fourth consideration is the specificity of association. When the same constellation of clinical and laboratory findings is inevitably associated with the same exposure, then this provides further evidence of the causality of the association. Finding that a particular supposed cause is not associated with any other clinical syndrome further strengthens this interpretation. Lack of specificity has been one of the arguments used against the causal relationship between HIV and AIDS, as the majority of diseases associated with AIDS in Africa, such as wasting syndrome, diarrhoeal diseases, and tuberculosis, were severe burdens prior to the HIV era. Scientists who support the HIV-AIDS causal relationship point to the marked change in the epidemiology of severe disease, as proof of HIV's role. High mortality rates due to these diseases now occur among HIV-infected young and middle-aged people, while death was formerly largely confined to the elderly and malnourished.

A clear dose-response relationship provides further support for a causal relationship. Although direct measurement of infecting dose is not usually feasible, an indirect indicator of the dose-response relationship may be the documented worsened AIDS prognosis when there is an increased quantity of virus in circulation.

Biological plausibility is an additional criterion but should be treated with caution. Most infectious agents have been associated with the disease they cause long before their
pathogenic mechanism has been completely elucidated. Work on the pathogenesis of HIV is ongoing.10

The “removal” criterion may provide additional support for a causal relationship i.e. where removal of the cause leads to elimination of the disease. This criterion depends upon an effective cure being available, which is not currently the case with AIDS. However proxy measures deserving consideration are the dramatic reductions in the incidence of AIDS and AIDS-related deaths in populations where potent anti-HIV combination therapies are widely available.11

None of these criteria can provide indisputable evidence for a cause-effect relationship between an exposure and a disease or in this case HIV and AIDS. They do however provide an explicit means of weighing up available evidence to judge causality, once chance, bias and confounding have been eliminated as the reason for the association. Proof of causality can have profound economic and social consequences. Effective corrective action to address the detected cause is demanded with often, considerable financial implications. Apportioning of blame may lead to calls for retribution, and the power bases of academics, politicians and other societal leaders may be threatened. However this should not deter explicit consideration of causality, as optimal patient management and effective interruption of disease transmission may depend on it.

References

David N. Durrheim MBChB, DTM&H, DCH, FACTM, MPH &T&M, MACTM
Associate Professor, School of Public Health and Tropical Medicine, James Cook University Australia & Consultant, Communicable Disease Control, Department of Health, MPUMALANGA

Gboyega A. Ogunbanjo MBBS, MFGP (SA), M Fam Med (MEDUNSA)
Associate Professor, School of Public Health and Tropical Medicine, James Cook University Australia

NB: The authors will welcome constructive comments on this article from readers.