HIV phylogenetics
Criminal convictions relying solely on this to establish transmission are unsafe

The recent flurry of criminal cases brought against people in the United Kingdom accused of infecting their sexual partner(s) with HIV has resulted in several convictions. This has caused concern among health professionals and community groups about the detrimental effect such cases may have on disclosure of HIV infection and uptake of voluntary HIV testing, which contrasts with the move to normalise HIV testing and clinical care. The potential negative effect of this on the public health programme to reduce transmission of HIV has been widely discussed in these pages and elsewhere.

Virological evidence, specifically HIV gene sequence data obtained from the defendant and complainant, has been used in these cases because a prerequisite for establishing criminal liability is that the
defendant caused the complainant’s infection. Because HIV-1, like other RNA viruses, evolves rapidly, the virus isolated from independently infected people is typically distinct. The extent of similarity between viruses from different people is associated with the likelihood of a common source of infection.

Within infected people, viral populations expand from a small initial population, and the viral genome can change during chronic infection by up to 1% each year. Thus, an onward transmission, if it occurred some time after initial infection, is likely to involve a strain distinguishable from the host viral strain. Viral phylogenetics, the scientific study of evolutionary relations between strains, has previously been used successfully to explore occupational and healthcare related transmissions in the context of detailed epidemiological information in cases in Florida, Baltimore, and Louisiana.

In recent criminal cases in the UK, attempts have been made to present evidence on the viral sequence as a conceptual framework akin to DNA fingerprinting (for which the odds of DNA samples from two different people being the same is typically lower than the population of the planet). In our view, this analogy is seriously misleading for several reasons, and when attempting to establish that transmission occurred between specific people virological evidence should be used with caution and only in conjunction with the clinical and epidemiological evidence.

Viral phylogenetics provides methods for assessing the relations between viruses from different people. A phylogeny is a hypothesis under which we can estimate the probability that viruses from two particular people have a recent common origin, but only in relation to other strains compared. The reliability of the conclusion depends both on the assumptions made in the statistical analysis and the data available for analysis.

Statistical tests (parametric, non-parametric, or Bayesian) are used to evaluate one such hypothesis against another, but if an inaccurate model is specified, or if the most relevant data are not included, these tests can, perhaps counterintuitively, give formal support to an incorrect conclusion. There are serious limitations on what can and cannot be inferred using phylogenetics alone and, in our view, a conviction that relies on such evidence to establish transmission is inherently unsafe. For these reasons, expertise should be sought before undertaking such analyses.

The greatest difficulty lies with the nature of the data. Identifying a linkage between viruses from two people on its own says nothing about the direction of transmission (who infected whom?) without allied information for the individuals concerned, and multiple specimens may be needed from before and after infection. Secondly, it is unlikely that all sexual contacts of all HIV infected people will be available for viral testing; indeed some may not be diagnosed. Thus, it is extremely difficult to distinguish a direct transmission between two people from a transmission from a third party to both, or from the first to the second person indirectly through a third. Thirdly, because HIV infected people can be coinfected or superinfected with genetically diverse strains, interpretation of the phylogeny is even more complex. Finally, similarities in two virus genomes may occur as a result of convergent or parallel evolution. An example of such is the independent development of drug resistance mutations, which can erroneously link people with no history of direct contact. Now that this is known, genetic positions subject to mutation associated with drug resistance can be excluded, but other sources of convergent evolution, such as those that might be produced by the virus evolving to escape the immune system, are not as well characterised.

We therefore advise caution when interpreting such data because the strength of any apparent linkage between viruses will never approach the degree of certainty generally expected of “DNA” data in a criminal court, which juries are more familiar with. Phylogenetic evidence—together with clinical and epidemiological evidence regarding likely duration of infection, sexual history, and other relevant factors—can provide support for linkage between cases but cannot prove transmission.

In this context, the only safe use of virus gene sequences is in circumstances where the genetic differences between viruses are sufficient to make linkage between two people doubtful (or in the case of different HIV-1 subtypes, highly unlikely). This is equivalent to the ability of blood grouping to establish that two samples come from different people (when the blood group differs), but not that they come from the same person (when the blood group is the same).

Despite the difficulty in determining linkage between specific individuals, phylogenetics can provide important new insights in investigations. A recent example is a study of the timing of HIV-1 infections among Libyan children in hospital, which showed that most infections occurred before the arrival of the accused medical workers in the country. Within the UK, new diagnoses comprise imported infections, many from sub-Saharan Africa, together with a growing number of infections acquired within the UK, overwhelmingly in men who have sex with men. The speed with which imported infections, from Eastern Europe as well as Africa, will lead to ongoing spread within the UK is unknown, as is the future mixing of viruses between different risk groups (those practising heterosexual or homosexual sex).

Molecular epidemiological approaches, allied to existing surveillance of HIV, will allow sensitive real time monitoring of such trends to be established, thus guiding targeted and cost effective public health interventions.

It will be important that sufficient checks and balances are in place to allow full use of such data for public health benefit, without concern that the underlying purpose for identifying possible viral genetic linkage between people will be to support criminal proceedings.

References are on bmj.com