THE ROYAL SOCIETY

doi 10.1098/rstb.2001.0877

The distribution of early acquired immune deficiency syndrome cases and conditions for the establishment of new epidemics

Daniel Low-Beer

Sidney Sussex College, University of Cambridge, Cambridge CB2 3HU, UK (dll0004@cam.ac.uk)

This article presents discussion on two issues: the distribution of early cases of acquired immune deficiency syndrome (AIDS), and the epidemiological conditions for the establishment of new diseases. Evidence from four presentations on early AIDS cases is discussed, together with issues of interpreting association and inferring causation. In the second section, the results of a stochastic epidemiological model using the 'real' geography of villages in northern Zaire is presented. This is used to investigate the conditions under which newly introduced infections with different characteristics are able to establish in human populations. It highlights additional spatial, temporal and behavioural conditions necessary for the persistence of introduced diseases, in addition to the condition that the basic reproductive rate $R_0 > 1$.

Keywords: early AIDS cases; association; causation; persistence of new diseases; stochastic geographical modelling

At this stage, it is timely to discuss two issues arising from the proceedings over the Discussion Meeting. First, several participants presented evidence on the distribution of early reported cases of acquired immune deficiency syndrome (AIDS). How have these associations been interpreted? Second, the more general point, that the introduction of new viruses is only the first chapter in determining whether they establish themselves in human populations. I want to look at some of the epidemiological constraints on this process.

In the map in figure 1, the argument we have heard on the distribution of reported AIDS cases is represented, as explicitly as possible, based on the data from Ed Hooper and the work of Sidharta Singh from Cornell University NY, USA. The sites of reported African AIDS cases up to 1980 (related to human immunodeficiency virus type 1 (HIV-1)) are shown as triangles. They are concentrated in ex-Zaire, but cases also occur in neighbouring Burundi, Rwanda, Zambia and Uganda. They suggest rural as well as urban epidemics. The sites are dispersed by quite large distances of up to 1000 miles.

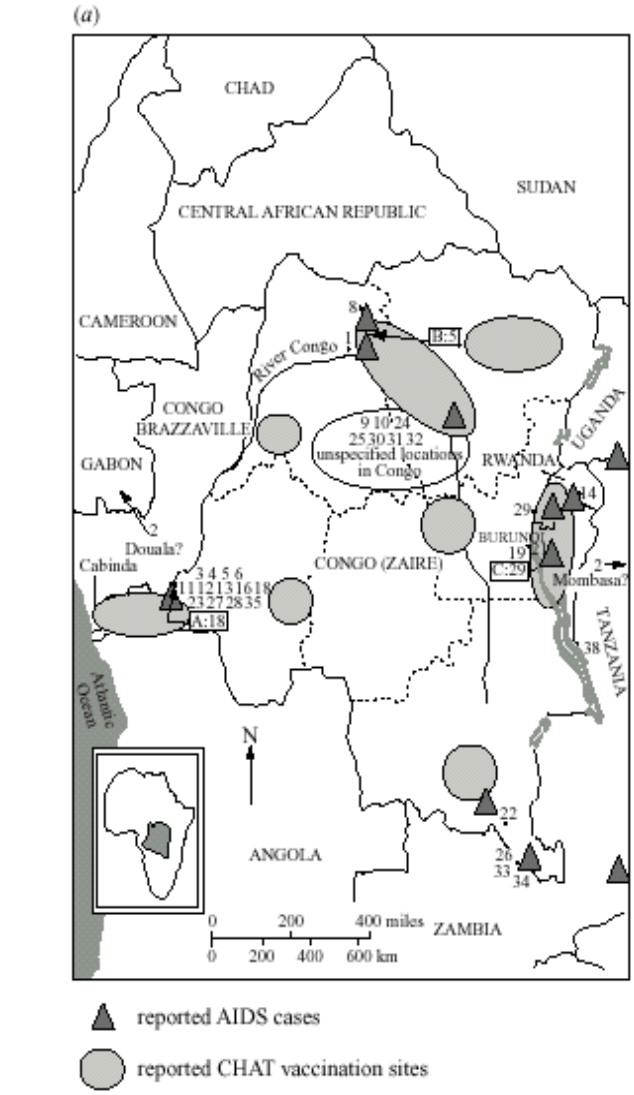
What has been suggested concerning these early AIDS cases? First, the oral polio vaccine (OPV) hypothesis has presented a geographical association with CHAT vaccination sites (denoted by the shaded ovular areas). Evidence has also been provided that reported African AIDS cases before 1979 are a mean distance of 40.4 miles from the vaccination sites, increasing to 200 miles by 1980 (figure 1*b*).

Stanley Plotkin stated that the Lubudi vaccination site in the south did not exist. Ed Hooper's evidence for its existence is from a newspaper article. Stanley Plotkin also provided evidence from four maps of the existence of vaccination sites to the west of Kinshasa, where there is no evidence of early AIDS cases. There is evidence of several other vaccination sites where there are no reported AIDS cases.

In understanding association, it has been stressed that we need to take into account confounding factors, and ascertainment bias. Kevin de Cock also emphasized that these reported cases are only a sample of possibly many thousands of cases circulating by the mid-1970s in ex-Zaire and elsewhere in Africa (reported cases are 'the tip of the iceberg of actual cases'). David Serwadda challenged the extent of ascertainment bias, suggesting that if there were large numbers of AIDS cases, they would have been noticed in Africa ('we never saw an iceberg in Lake Victoria'). He noted that early AIDS cases were identified by communities in Uganda and described as Slim. However, it must be emphasized that we are dealing with reported cases, which provide positive but not negative evidence of occurrence. We need to look just as carefully at evidence for the absence as for the presence of AIDS cases in comparison areas (and tested sera that are HIV negative); 24 800 sera collected up to 1980 have been HIV tested from 28 countries in Africa, five times the evidence elsewhere, according to Smallman-Raynor *et al.* (1992). At the same time, it may be unlikely that there are large numbers of AIDS cases before the mid-1970s; rather, we are dealing with low-level chains of infection.

It is important to stress two geographical principles in looking at these data, even after we have assessed issues of bias.

First, association is not sufficient for causation: therefore discussion has been rightly focused on the existence of a biological mechanism. Were chimpanzees used in OPV trials? Were their kidneys infected with simian immunodeficiency virus (SIV)? Is this likely to pass on



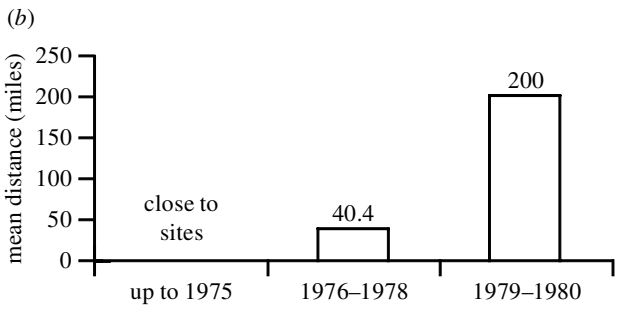


Figure 1. (*a*) Evidence presented of early reported AIDS cases (HIV-1) and association with CHAT vaccination sites. (*b*) Mean distance, in miles, of reported AIDS cases from CHAT vaccination sites in different time-periods, up to 1975, 1976–1978 and 1979–1980. Source of data: Hooper (1999, pp. 742–747).

infection to significant numbers of humans if administered orally? Although we cannot dismiss evidence of association and many epidemiological hypotheses are developed by critical analysis at this level, we need to understand better the context of early HIV spread, and move the argument to assess the plausibility of biological mechanisms.

Second, several processes can result in the same geographical distribution. When we look at the vaccination sites, there are transport routes nearby. Stanley Plotkin has suggested that the early reported AIDS cases were sexually transmitted and resulted from chains of infection from outside the areas where they were found. These explanations, in particular linking reported AIDS cases and HIV prevalence in the early 1980s to diffusion along transport routes, certainly need to be formulated explicitly and discussed for the validity of the various hypotheses of the origins of AIDS.

The map in figure 2 shows the extensive transport network in Africa, and the sites of the early reported AIDS cases. Clearly in most instances we are not dealing with isolated sites, for the reported AIDS cases or OPV immunization. The pattern of distribution of early AIDS cases may suggest locations for the origins, or of diffusion processes from outside. These could both show similar distributions of early AIDS cases, though rarely with the same probability. A number of questions were raised concerning the relationship to transport routes. Why are the early AIDS cases over 1000 miles from the origin for 'natural transfer' suggested by present evidence of SIV among primates? If AIDS cases spread along the transport routes, why are early AIDS cases not reported elsewhere to the south, west, and east, and significant urban epidemics initiated earlier? Other cases may have existed, but were not reported. But two pieces of further evidence would be important: evidence of early AIDS cases elsewhere (and closer analysis of non-occurrence) and variations in HIV prevalence in the early 1980s; or evidence of primate hosts for 'natural transfer' in Central or East Africa, an issue which Beatrice Hahn is pursuing.

This brings me to the second issue: introduction events are only the preface to the establishment of new diseases in human populations. At the University of Cambridge, we have developed stochastic epidemiological models based on the 'real' geography of the villages around Yambuku in northern Zaire (figure 3*a*). These villages show some of the first signs of HIV infection and previously of Ebola. We investigated under what conditions newly introduced infections, with different characteristics, become established in this village network.

The epidemiological problem in these villages is stochastic, as with 1% HIV prevalence we are dealing with small numbers of infections, prone to random extinction. Traditional deterministic epidemic models show that infections persist if their basic reproductive rate is greater than one. Furthermore, HIV infection at low levels can be maintained for several decades, given a particular pattern of infection. However, epidemic simulations in our stochastic model (figure 3*b*) reveal some problems. Low-level epidemics are very prone to extinction over this period of low-level infection. This occurs in most simulations. However, stochastic models also allow us to analyse additional characteristics of epidemics that persist.

In epidemics that persist, we find a 'star-like' geography of infection in a cell of interacting villages, which is established very early on. This reduces the probability of random extinction. This pattern can also occur over time with a 'pulse' pattern of multiple introductions. Emerging epidemics behave more like individual realizations of stochastic simulations (with many others dying out) than

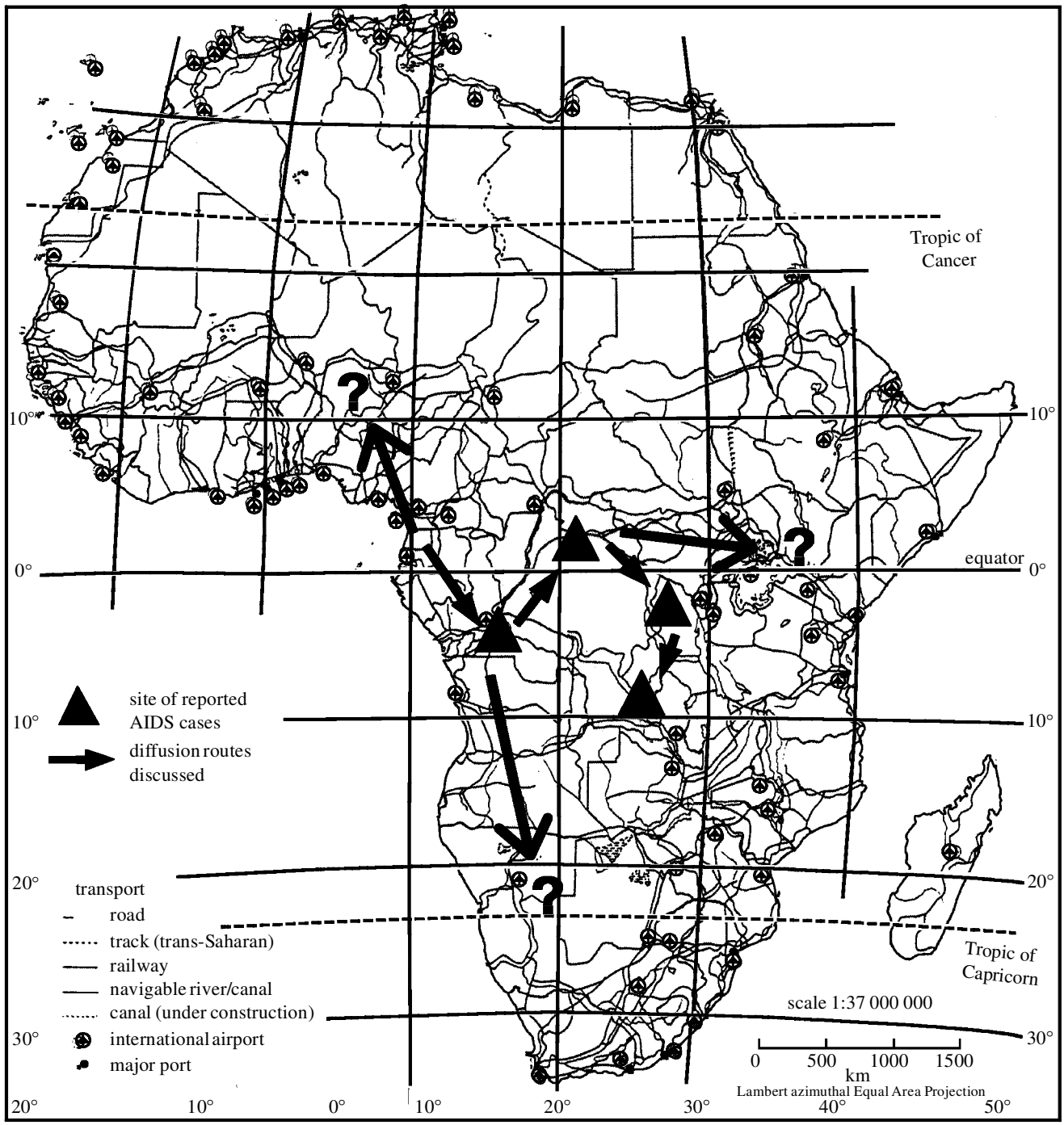


Figure 2. Major transport routes in Africa, sites of early reported AIDS cases (triangles), and diffusion routes discussed (arrows). Arrows show the diffusion routes raised in discussions, and questions as to why reported AIDS cases are concentrated on a restricted section of the transport route and not elsewhere. Map source: *The atlas of Africa*, p. 68.

either a mean of stochastic simulations or a deterministic model.

If rural areas are the sources of HIV epidemics, the introduction events must generate sufficient contact to establish this cell of interacting village infections very early on. We are therefore less likely to be discussing single natural transfer events (the mythical cut-hunter) or a few OPV infections, but sustained contact geographically or over time.

The stochastic model also suggests additional behavioural constraints to persistence. Increases in risk behaviour over time (or lower risk historically) compound the probability of earlier stochastic extinction, and decreases that of later stability. In these rural areas, risk behaviour decreasing over time is most consistent with low-level epidemics (for example, due to increasing outmigration of young adults). Introducing migration into the geographical network helps explain low, stable HIV prevalence in some rural areas (and resulting urban epidemics) and less stable, increasing rural epidemics in areas of return migration (for example, in Uganda).

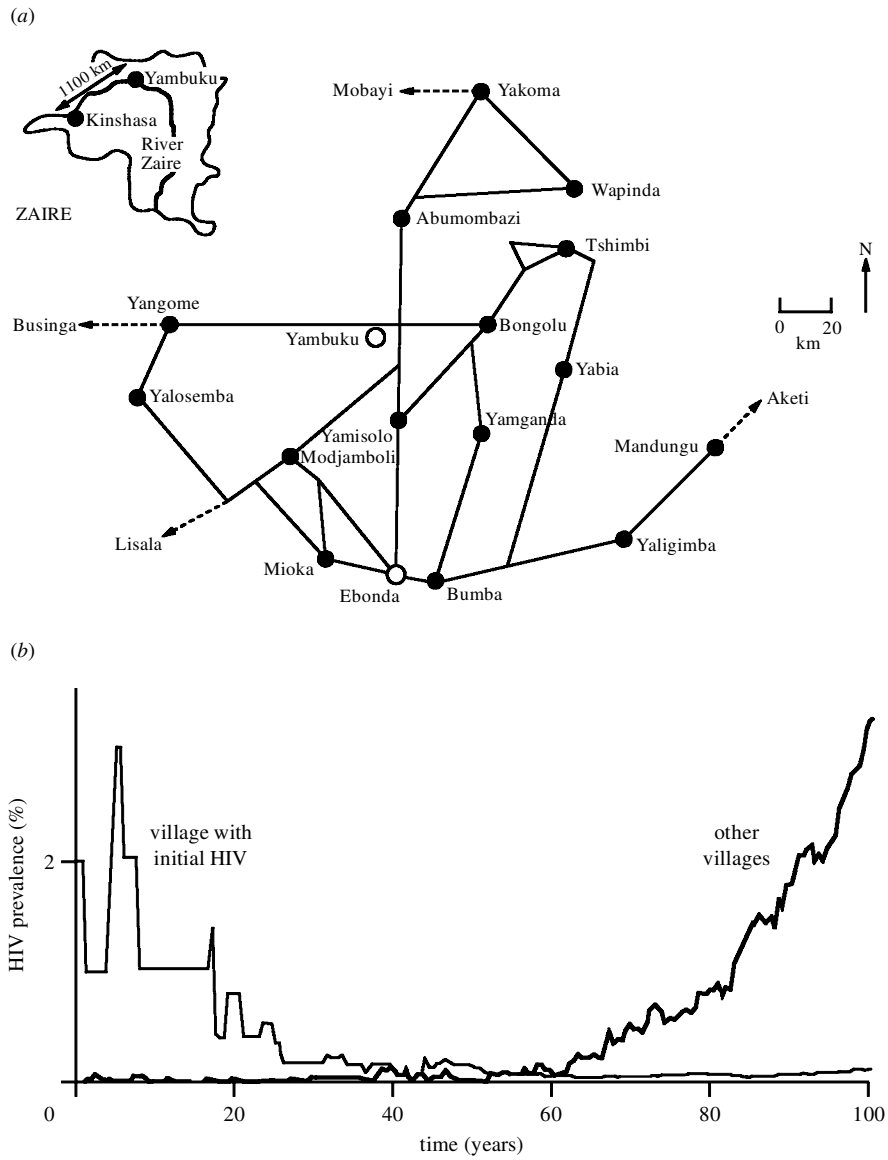


Figure 3. (*a*) The village network in northern ex-Zaire around Yambuku (in the 1970s) used as the basis of epidemiological modelling. (*b*) A stochastic epidemic simulation showing HIV infection maintained at low levels in the village network for several decades (based on a mixing matrix of intra- and inter-village contact combining epidemiological and geographical parameters). HIV prevalence for the village where HIV is introduced and other villages in the network, by year.

A possibility nevertheless remains that the driving cells of infection are outside these rural villages (maybe in particular groups in cities), and the 'smouldering' rural infections we see, are successive secondary chains of infection that die out. The approach has highlighted additional spatial and behavioural constraints to the emergence of new diseases, which should be discussed in assessing the validity of introduction events.

The stochastic modelling also illustrates a final more general point (figure 4). Whatever the origins of HIV, there is a stochastic hurdle to the persistence of single introductions of low-level infections: multiple introductions are a particular concern, and may expose populations to a set of infections that may not normally persist in humans.

Relatively infectious diseases like Ebola or influenza have a high probability of persistence (shown on the *y*-axis) even with a few initial infections (on the *x*-axis). However, there is a range of diseases with lower

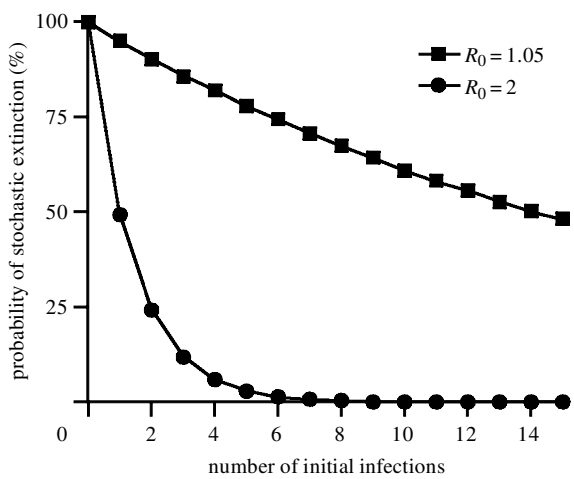


Figure 4. Probability of stochastic extinction of an epidemic by number of initial infections for infectious agents with $R_0 = 1.05$ and $R_0 = 2$.

infectivity (but $R_0 > 1$), which are only likely to establish in human populations after multiple introductions (i.e. introduced further along the *x*-axis). This is significant for both 'natural transfer' and OPV introductions, which result in low-level epidemics.

In conclusion, the association between early reported AIDS cases and CHAT vaccination sites may not simply be dismissed, but needs to be looked at together with sources of bias and evidence for the non-occurrence of early cases elsewhere. However, it is now important to go beyond association and, in analysing hypotheses of origin, look for biological causation and characteristics that fit the epidemiological constraints of establishment in a human population.

I would like to acknowledge contributions from Sidharta Singh, Cornell University, USA, in the first section of this presentation, and Professor Andrew Cliff, Department of Geography, University of Cambridge, UK, in the second section, and Dr Thomas Low-Beer for comments.

REFERENCES

Editions Jeunes Afrique 1970 *The atlas of Africa*, 1st edn. Paris: Editions Jeune Afrique.

Hooper, E. 1999 *The river: a journey back to the source of HIV and AIDS*. London: Penguin.

Smallman-Raynor, M., Cliff, A. & Haggett, P. 1992 *Atlas of AIDS*. Oxford, UK: Blackwell.