Surrogate markers in HIV disease

Tim Peto

Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK

The use of surrogate markers in HIV disease is an attractive method of assessing the efficacy of new treatments more quickly than by using clinical end-points. The characteristics of an ideal surrogate marker and the theoretical dangers of extrapolating properties from one class of drug to another are described. These characteristics are compared with the use of the CD4 lymphocyte count, which so far has been the most widely studied. Results from 14 randomized controlled trials of nucleoside analogues are used to compare the comparative changes of CD4 counts with the differential rates of progression to AIDS and differences in survival. There was some correlation between CD4 count changes and development of AIDS, particularly in the short term trials. In contrast, there was little correlation between CD4 counts and overall survival. Comparative studies between clinical end-points and quantitative measures of plasma viraemia have not yet been completed. In conclusion, no surrogate marker has yet been shown to be useful in predicting the efficacy of anti-HIV treatment. Until surrogate markers are validated against the results from long term clinical trials, they should only be used to screen new drugs warranting further study rather than to draw conclusions on the clinical efficacy of new treatments.

Introduction

There are many different anti-HIV drugs being developed and there is a great wish to assess new treatments and complex drug combinations as quickly as possible. However, the slow natural history of HIV infection means that controlled trials using mortality as an end-point inevitably take several years to complete. Even trials on symptomatic patients, in whom death occurs more rapidly, require many thousands of patient-years of observation. It is therefore highly desirable that techniques should be developed which can assess more quickly the effectiveness of drugs. In order to achieve this, surrogate markers, or 'intermediate end-points' have been adopted in the hope that they are reliable and effective. CD4 lymphocyte counts have been widely used as an obvious choice as a surrogate marker. For example, in 1991, an advisory committee to the US Food and Drugs Administration recommended that didanosine be approved for use in persons on the basis of changes in the CD4 lymphocyte counts of subjects within 6 months of the initiation of therapy rather than on clinical evidence of benefit. Unfortunately, the results of two large scale randomized trials of antiretroviral therapy have shown that the changes in CD4 counts do not extrapolate in a simple fashion to changes in disease progression or mortality (Darbyshire & Aboulker, 1992; Concorde
Coordinating Committee, 1994). In this article the ideal properties of surrogate markers are described and compared with the currently available HIV surrogate markers.

**Ideal properties of surrogate marker**

Surrogate markers have been used to assess prognoses, monitor therapy and assess treatment efficacy. The ideal property of a marker will depend on its use.

**Prognostic markers**

There is a clinical need to stage HIV disease and to determine the risk of disease progression. Results from observational studies are sufficient to validate the use of markers as prognostic indicators. However, their clinical utility does not critically depend on the marker being causally related to the disease. For instance, in renal disease, plasma creatinine is an excellent prognostic marker for renal failure but it is well known that it is not the cause of renal failure.

In HIV disease, CD4 lymphocyte counts are useful for determining prognosis and susceptibility to opportunistic infections (Phillips et al., 1991; Saah et al., 1992). However, there is much epidemiological data showing that CD4 lymphocyte counts do not alone account for all the variability in the risk of disease progression (Choi et al., 1993; Lagakos, 1993). Table I shows other clinical and laboratory markers which have been shown to be related to disease progression. Apart from the CD4 count, which is believed to be directly related to the total body CD4 mass, other measures of disease progression include markers of immune activation, immune dysfunction and clinical evidence of disease progression (Lange, de Wolf & Goudsmit, 1989; Fahey et al., 1990; Jacobson et al., 1991; Lafeuillade et al., 1994; Tsoukas & Bernard, 1994).

**Monitoring of treatment**

There is often a therapeutic need to determine the pharmacological efficacy of treatment. Simple examples include the use of monitoring drug concentrations in blood or determining the immediate effects of treatment. Examples in other diseases include the use of blood glucose and circulating haemoglobin A1C concentrations to monitor insulin treatment of diabetes mellitus, and blood pressure in the treatment of hypertension. In many cases, the causal role of the surrogate marker in the evolution of disease may be in doubt, but the marker can be useful in demonstrating indirectly a pharmacological action of treatment. Markers for drug resistance (for example, antibiotic resistance) can be used to predict changes in the immediate pharmacological effect of treatment. The interpretation of these surrogate markers depends on an understanding of the pharmacological action of the drug.

Most anti-HIV drugs have been selected for clinical use by their effectiveness in inhibiting HIV replication in vitro. Changes in plasma viraemia, with treatment, can therefore be considered a direct or immediate effect of treatment. Isolation of drug resistant viruses are being used to predict a failure of the drug to change viral load and therefore by extrapolation to predict clinical efficacy.
Surrogate markers in HIV disease

Comparison of different treatments

Surrogate markers can be used to determine the clinical efficacy of a new treatment. This is the most stringent requirement for a surrogate marker. It has been suggested that for a surrogate marker to be valid "... a test of the null hypothesis of no relationship [of the marker] to the treatment groups must also be a valid test of the corresponding null hypothesis based on the true endpoint." (Prentice, 1989). For a surrogate marker to satisfy this, changes over time must correlate with the risk of progression and any effect of treatment on the risk of clinical progression must be explainable and predictable by the effect of treatment on the markers. For example, CD4 lymphocyte counts are a valid marker only if the whole treatment effect of zidovudine is reflected by changes in the CD4 counts. Unfortunately observational studies have failed to show this (Choi et al., 1993; Lagakos, 1993).

Even this stringent requirement of a surrogate marker has its limitations. It is also important to have data on the size of a treatment benefit in order to determine whether the benefit outweighs its toxicity and cost. Such a quantitative relationship between changes in the surrogate marker and the clinical outcome is difficult to predict. A common fallacy is to assume that the relationship between a prognostic marker and outcome, obtained from natural history studies, can be used to predict the effect of a treatment induced change in the marker.

The different possible relationships which might be found between a change in a surrogate marker and the clinical outcome is represented schematically in Figure 1. Each relationship represents the response of the same marker to different classes of drugs. The horizontal axis represents the effect of treatment on the value of the surrogate marker while the vertical axis represents any improvement (or worsening) in clinical outcome. In the diagram some responses are compatible with a 'Prentice' marker (Prentice, 1989). For example, drug 'O' is ineffective with no effect on the marker nor on outcome. Drug 'I' produces a dose-response curve of an ideal surrogate marker in which there is a continuous and predictable quantitative relationship between the change in the marker and clinical effect. Drug 'H' produces a large change in the marker

![Figure 1. Schematic diagram comparing the results of different classes of drugs (A–I); I being the 'ideal' response. The abscissa represents the differential effect on the surrogate marker and the ordinate the differential effect on clinical outcome.](https://academic.oup.com/jac/article-abstract/37/suppl_B/161/740752/740752)
with little benefit, while drug 'G' is clinically effective but only has a marginal impact on the surrogate marker. Drug 'B' worsens both outcome and the marker. The other drugs (A,C,E,D) do not follow the expected surrogate-clinical relationship. Drugs E and C affect clinical outcome without changing the marker. Pathophysiological explanations are not difficult to conceive: for instance, a clinically effective drug (E) could affect the distribution of lymphocytes (or virus) between the circulation and other tissues. An inactive drug (C) could be dangerously toxic. Other clinically ineffective drugs could simply affect the marker as an 'epiphenomenon'.

Drugs may have the opposite effect than expected (for example, A and D). Drug D looks attractive as judged by the surrogate response but is clinically deleterious. This pattern of response has often, and dramatically, been shown in other diseases. In particular, in heart disease, flecainide was shown to prevent arhythmias which are prognostically a dangerous sign but perversely was shown in a large randomised study to increase mortality (Cardiac Arrhythmia Suppression Trial, 1989). Drug A is an important class of drug which is likely to have a novel action producing clinical benefit in a manner that worsens the prognostic marker. It is not implausible that a new 'miracle drug' which kills all HIV infected cells would cure the patient but might produce a transient although profound fall in circulating CD4 cells and an intense rise in the measured plasma viral load. Such a paradoxical effect on prognostic markers is illustrated by the effect on haemoglobin and white blood counts of chemotherapy for the treatment of leukaemias. The response of a surrogate marker to a new class of anti-HIV therapy should therefore be interpreted with caution until its relationship with the outcome has been determined in clinical trials.

Choice of surrogate markers

A rationale choice of surrogate markers is not easy. It is tempting to use a theory of the patho-physiological process underlying the evolution of the disease (Figure 2). However, epidemiological studies identifying prognostic factors associated with disease progression provide an alternative method to screen for putative surrogate markers. Demonstration that these factors are causally associated with progression disease is difficult without evidence from randomized intervention studies. Surrogate markers can be divided into two main types: markers of disease progression (Table I) and markers of disease activity (Table II).

The surrogate marker hypothesis

\[
\text{viral susceptibility} \xrightarrow{\text{drug treatment}} \text{plasma viraemia} \xrightarrow{} \text{CD4 lymphocyte count} \xrightarrow{} \text{clinical signs of progression} \xrightarrow{} \text{death}
\]

Figure 2. Diagramatic representation of the causal assumptions underlying the interpretation of surrogate marker data.
Markers of disease activity

The ideal markers for disease activity should directly measure the pathological process which are causally responsible for clinical disease. Table II shows the surrogate markers that have been used as an index of disease activity.

Circulating viraemia. The advantages of measures of viral load are that they offer a pathophysiologically plausible index of disease activity and that they are easy to quantify (Mellors et al., 1995). They are now commonly used to screen the clinical activity of new drug combinations (Katzenstein et al., 1994; Lafeuillade, et al., 1994; Molina et al., 1994). Repeated measurements can be easily performed on each patient and changes in viral load accurately followed over weeks and months. Therefore, only a small number of patients need to be studied to obtain a repeatable result.

The disadvantages, however, of the assays are fundamental. Firstly, it is unclear whether circulating plasma viral products (RNA, viral protein or intact virions) all give internally consistent results (Bieniasz et al., 1993). Secondly, it is unclear whether plasma viral load is the main measure of HIV disease activity. Recent work has suggested that the main reservoir of HIV disease is in the lymph nodes (Embretson et al., 1993; Graziosi et al., 1993; Pantaleo et al., 1993). Although observational studies suggest that plasma viral load gradually increases with disease progression, the importance of circulating viral load is unclear. It is possible that most viruses reside in tissue cells and only use circulatory cells to traffic between tissues. It is also possible that cell-free viral load represents the products of cell death and therefore might be expected to rise in the face of effective anti-HIV therapy. Thirdly, the quantitative relationship between changes in viral load and drug efficacy is completely unknown. There is a reasonable belief that the more viral load is suppressed and the more prolonged the trough, the better. However, this assumption has never been tested. Recent studies on the kinetics of viral replication suggest that viral turnover is very marked (Ho et al., 1995; Wei et al., 1995). This makes any inference from viral load measurements to total body viral replication difficult. For instance, drugs which increase

<table>
<thead>
<tr>
<th>Table I. Examples of markers of disease progression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measures of CD4 body mass</td>
</tr>
<tr>
<td>Landmark levels of CD4 count, %CD4 count, CD8, CD4/CD8 ratios</td>
</tr>
<tr>
<td>Immune activation</td>
</tr>
<tr>
<td>beta 2 microglobulin</td>
</tr>
<tr>
<td>neopterin</td>
</tr>
<tr>
<td>Immune dysfunction</td>
</tr>
<tr>
<td>cutaneous anergy</td>
</tr>
<tr>
<td>Clinical events as surrogate markers for disease progression</td>
</tr>
<tr>
<td>symptoms of ARC</td>
</tr>
<tr>
<td>first AIDS defining symptoms</td>
</tr>
<tr>
<td>incidence of recurrent AIDS events</td>
</tr>
<tr>
<td>TB</td>
</tr>
<tr>
<td>weight</td>
</tr>
<tr>
<td>Karnofsky score</td>
</tr>
<tr>
<td>Virological markers of immune dysfunction</td>
</tr>
<tr>
<td>syncytium-inducing HIV-phenotypes</td>
</tr>
<tr>
<td>Plasma HIV RNA load</td>
</tr>
</tbody>
</table>
the clearance of virions from the plasma will lead to an observed fall in virion concentration, which in turn may lead to a false conclusion that total body viral replication has been suppressed.

Lastly, the relationship between drug efficacy and changes in viral load may differ between different classes of antiviral drugs. Although a rational relationship may exist between viral load changes and efficacy for nucleoside analogue reverse transcriptase inhibitors, the same relationship may not hold for other reverse transcriptase inhibitors, let alone protease inhibitors. Clearly, the expected short term effect on circulating viral load of effective gene therapy strategies or other novel treatments (for example, the transfer of HIV-1 specific cytotoxic lymphocytes (Koenig et al., 1995)), is completely unclear. The interpretation, therefore, of changes in the markers of disease activity with novel treatments is difficult and may lead to major errors. Current large scale randomised clinical trials of combination treatment (DELTA and ACTG 175) are designed to compare changes in viral load, as judged by these new virological assays, with clinical outcome.

Changes in CD4 counts. Zidovudine and other nucleoside analogues have been shown repeatedly to increase CD4 counts transiently for up to 6 months. The validity of the CD4 count has been tested against the gold standard of randomized controlled trials measuring clinical disease progression and survival with zidovudine and didanosine. A large number of trials have now been completed but have measured CD4 responses in slightly different ways. Also, on an intent-to-treat basis the length of follow up, varies making comparison between trials difficult. Nevertheless, it is interesting to compare the different results. Fleming (1994) has collected the results from 14 randomized trials of AZT vs placebo. In some cases, Fleming (1994) re-analysed the trials using unpublished sources in order to provide an unbiased intent-to-treat analysis of the rates of progression to AIDS and death. For convenience, these results are presented graphically, but no allowance has been made for the small size of some of the studies. Figure 3 shows the relationship between the 'average' difference in the absolute CD4 response between treatments and the relative rates of progression to AIDS during each of 12 studies. In spite of the small size of some of the trials, different baseline CD4

![Figure 3. Comparison of differential changes in CD4 counts and relative risks of progression to AIDS or death in 12 randomized trials of nucleoside reverse transcriptase inhibitors.](https://academic.oup.com/jac/article-abstract/37/suppl_B/161/740752/1041777074702)
Surrogate markers in HIV disease

Figure 4. Comparison of differential changes in CD4 counts and relative risks of progression to death in 12 randomized trials of nucleoside reverse transcriptase inhibitors.

counts at randomization, different durations of follow-up and different definitions of a CD4 response, a reasonable relationship appears to be shown demonstrating that effective nucleoside analogues need to increase CD4 counts by more than 30–40 to produce a change in the incidence in AIDS. In contrast, Figure 4 shows that the same CD4 changes do not predict improved survival.

The discrepancy between progression rates of AIDS and mortality is probably best explained by the different durations of treatment. Zidovudine treatment appears to delay the progression of disease and mortality for several months, but the effect disappears after prolonged treatment. This can be correlated by the development of virus resistance to zidovudine. It is therefore clear that it is difficult to infer a long term impact on clinical disease from short term changes in surrogate markers.

Markers of disease progression

Markers for disease progression have been used as a surrogate for clinical events. Possible surrogate markers are listed in Table I. They have been chosen because observational studies have shown them to be good prognostic markers for survival.

The most widely used surrogate marker for disease progression is the absolute CD4 count. It is a good prognostic marker and is easy to measure in large numbers of patients. It is likely that CD4 depletion is causally related to the incidence of HIV-related opportunistic infections and tumours. Some clinicians prefer to make treatment decisions with anti-HIV drugs with the intent of keeping CD4 counts high for as long as possible, in the belief that improving the prognostic marker will improve the prognosis. Unfortunately, this assumption has not been validated. Indeed some observational studies suggest that the use of zidovudine is a confounding factor in this relationship.

Nevertheless, some large scale randomized controlled trials have used CD4 counts as end-points in clinical trials (for example, Cooper et al., 1993). They are sometimes hidden because they are part of a definition for clinical end-points. The advantages
of using CD4 end-points are clear. Trials can be much shorter and fewer patients need to be randomized. An interesting ethical advantage of these shorter trials is that by the end of the trial period, most individuals randomized within the trial will still be sufficiently well to be able to benefit from open treatment with the ‘better’ arm of the study. However, the result of the Concorde study (Concorde Coordinating Committee, 1994) has very clearly demonstrated that the incorporation of CD4 counts in the definition of ‘clinical end-points’ produced an apparent advantage for the immediate use of zidovudine compared with its deferred use. This advantage disappeared when the end-points of progression to ‘AIDS or death’ or overall mortality were measured.

Furthermore, it is also not clear whether clinical signs of progression are themselves a good surrogate for overall survival. Again the Concorde study showed some suggestion, though not statistically significant, that there was a discrepancy between the rates of progression to ARC, AIDS and death. Data from 12 trials, already described above (Fleming, 1994), have been used to compare the relative rates of progression to ‘AIDS or death’ with the relative mortality. Figure 5 shows that there is little evidence of a useful relationship.

![Figure 5. Comparison of relative risks of progression to AIDS and risk of death in 12 randomized trials of nucleoside reverse transcriptase inhibitors.](https://academic.oup.com/jac/article-abstract/37/suppl_B/161/740752/161740752)
Surrogate markers in HIV disease

Conclusions
At present there is no convincing evidence that the current surrogate markers can be reliably used to predict the clinical efficacy of new treatments. Indeed proper validation will probably need to await the arrival of much more effective clinical treatments. Meanwhile, surrogate marker responses should only be used in the early assessment of new drugs, in order to help the selection of new drug regimens for large scale testing in long term clinical trials.

References


