Virologic and Regimen Termination Surrogate End Points in AIDS Clinical Trials

Peter B. Gilbert, PhD
Victor DeGruttola, DSc
Scott M. Hammer, MD
Daniel R. Kuritzkes, MD

Suppression of plasma human immunodeficiency virus (HIV) RNA levels has been widely accepted as an appropriate surrogate end point for HIV disease progression, and it is currently used as the primary end point to determine efficacy in many antiretroviral trials. However, this end point does not always measure other important effects of treatment, such as inducement of multidrug resistance, which depletes future therapy options, and toxic effects. An alternative that directly factors in these treatment costs is a composite *regimen termination end point*, defined as a protocol-determined change in regimen due to either virologic failure or treatment-related toxic effects. Pros and cons for using purely virologic vs various composite primary end points are discussed. Conclusions include (1) a trial's clinical objective guides the choice of primary end point, (2) a purely virologic end point is often preferable, (3) it may be important to analyze both end point types in interpreting study results, and (4) long-term clinical outcome studies are needed for identifying the most predictive surrogate end points.

JAMA. 2001;285:777-784 www.jama.com

A CRITICAL STEP IN THE DESIGN of clinical trials to evaluate the efficacy of anti–human immunodeficiency virus (HIV) therapies or to compare treatment strategies is the selection of the appropriate primary study end point. A well-designed phase 3 trial definitively assesses the effects of treatment on the chosen primary end point, thereby defining the role of the therapies or strategies in the clinical management of disease. Because of the rapidly evolving scientific understanding of HIV infection and its therapeutic management, it is important to continually reevaluate primary end points. The purpose of this article is to describe some limitations of primary end points in current use; to propose principles for choosing between end points that measure biological activity alone and composite end points that directly factor in treatment costs, such as resistance or toxic effects; and to identify the kinds of studies needed to provide objective criteria for end point selection.

Historically, primary end points based on clinical events such as acquired immunodeficiency syndrome (AIDS)–defining illnesses and death carried the greatest weight in guiding clinical practice because observed differences in these end points signify tangible differences in treatment benefit. However, as therapy improved, the low rate of disease progression made it impractical to use clinical events as primary end points in trials of short duration, especially in treatment-naive groups. Almost all antiretroviral trials now use a biological marker (eg, CD4 cell count or plasma HIV RNA level) as the primary end point. The ability of any given trial to answer the posed clinical question depends on whether the marker end point is indeed a surrogate for the clinical outcome of interest.

A marker end point is a good surrogate for a clinical end point if observed treatment effects on the marker reliably predict treatment effects on the clinical end point. For this condition to hold, the biological marker (1) must be correlated with the clinical outcome, and (2) must fully capture the effect of treatment on the clinical end point.[1] The first criterion is easily verified and often holds, but the second is difficult to validate and generally fails.[2-4] The CD4 cell count is an example of a biological marker known to be prognostic for risk of opportunistic infections and death and was commonly used in AIDS clinical trials as an efficacy end point, but ultimately it was found to have limited predictiveness for progression to AIDS and death.[5-9] In more recent years, the level of plasma HIV RNA has been shown to be a better prognostic marker for clinical progression than CD4 cell count in most[10-15] but not all[16,17] studies. Potent antiretroviral therapy can suppress plasma HIV levels below as-

Author Affiliations: Department of Biostatistics, Harvard School of Public Health, Boston, Mass (Drs Gilbert and DeGruttola); Division of Infectious Diseases, Columbia Presbyterian Center, New York, NY (Dr Hammer); and University of Colorado Health Sciences Center, Denver (Dr Kuritzkes).
Financial Disclosures: Dr DeGruttola was an ad hoc consultant to or recipient of honoraria from Bristol-Myers Squibb, Gilead Sciences, Glaxo Wellcome, Tibatec, Virco, and Visible Genetics. Dr Hammer was an ad hoc consultant to, site investigator for, or recipient of honoraria from BioChem Pharma, Bristol-Myers Squibb, Boehringer Ingelheim, Gilead Sciences, Glaxo Wellcome, and Triangle Pharmaceuticals. Dr Kuritzkes received research grants and funding, honoraria, lecture sponsorships, and assay kits or reagents from or had consultancies with the National Institutes of Health (AIDS Clinical Trials Group), Abbott, Agouron, Bristol-Myers Squibb, Dupont Pharma, Glaxo Wellcome, LabCorp, Merck, Roche, Triangle, Trimeris, Virco, ViroLogic, and Visible Genetics.
Corresponding Author and Reprints: Peter B. Gilbert, PhD, Department of Biostatistics, Harvard School of Public Health, 655 Huntington Ave, Boston, MA 02115 (e-mail: pgilbert@hsph.harvard.edu).

Definitions of Primary End Point Types

Purely Virologic End Point. Time from randomization to virologic failure, with virologic failure defined by a confirmed rise in plasma human immunodeficiency virus (HIV) RNA levels above a threshold such as 200 copies/mL. Virologic failure may also include early virologic failure events such as lack of initial virologic response within 4 to 12 weeks or early virologic relapse, defined by a confirmed 1 $\log_{10}$ (10-fold) increase above a subject's lowest HIV RNA measurement (nadir) or by a rise above an absolute threshold.

Regimen Termination End Point. (1) Time from randomization to earliest event of virologic failure, permanent study treatment discontinuation, acquired immunodeficiency syndrome–defining event, and death. All treatment discontinuation events are counted as end points, regardless of the reason for discontinuation. (2) Time from randomization to earliest event of virologic failure and permanent study treatment discontinuation due to protocol-defined toxic effects. Only the subset of treatment discontinuation events confirmed to be due to protocol-defined toxic effects are counted as end points.

Mechanism for Handling Study Dropout

Each end point above must use a convention for classifying study dropout. Two conventions are commonly used:

Dropouts as Censored. Subjects who withdraw from the study prior to meeting an end-point definition are censored at the time of last contact, ie, have a failure date known only to exceed the date of last contact, and are considered to be successfully treated at that time.

Dropouts as Failures. Subjects who withdraw from the study are considered to have reached an end point on the date of last contact.

say detection limits in many patients.[18-26] Durable virologic suppression confers a significant reduction in AIDS-defining events and death[20-29] and slows or prevents the development of drug resistance.[30] This has led many investigators to accept end points based on plasma HIV RNA levels as primary end points in antiretroviral trials, although the surrogacy of virologic end points for clinical end points has not been fully validated. The reductions in rates of AIDS morbidity and mortality in the developed world resulting from the widespread use of drugs found to suppress viral replication in clinical trials[23-27] imply that use of virologic end points has been productive in the short term. However, the comparison of virologic activity across regimens gives an incomplete picture of the clinical differences. A purely virologic end point may not always be an adequate surrogate or even the best available surrogate for clinical end points.

Treatment-related toxic effects,[31] adherence difficulties,[32-34] and drug resistance[35] may make it necessary to use several regimens in sequence to durably control HIV replication. These complications prompt consideration of an alternative to a purely virologic primary end point in clinical trials of antiretroviral drugs (see BOX). One alternative is designated as the regimen termination end point (ie, the treatment failure point at which a regimen's benefit for a patient is "used up," possibly due to 1 or more factors). In practice, the *regimen termination end point* is defined as first occurrence of any protocol-specified event that leads to cessation of the assigned regimen. This does not necessarily imply that all regimen components have been expended. If resistance to or toxic effects from only 1 agent within the regimen led to its termination, other agents within the regimen might still be useful in subsequent regimens. Examples of AIDS trials that have used a regimen termination end point are given in the TABLE.[36-46]

Experience with surrogate markers in other disease areas provides useful lessons for end point selection in AIDS trials. Cancer researchers have long debated the use of surrogate biological marker end points (eg, tumor shrinkage) and composite treatment failure end points that include treatment discontinuation.[47,48] Treatment effects on surrogate end points have given false-positive or false-negative predictions of treatment effects on clinical outcomes in trials involving a variety of diseases. Examples include trials of arrhythmia-suppressing drugs[49,50] and a trial of interferon gamma for treatment of chronic granulomatous disease.[51] Relevant lessons from these examples are that clinical outcome studies are essential for defining the appropriate use of surrogate markers, and the intent-to-treat (ITT) principle is the best available analytic technique for handling inability or unwillingness to comply with treatment.[52,53] Furthermore, reporting analyses of both biological marker and treatment failure end points may aid in the interpretation and clinical application of the primary study result.[47]

Given the observations described above, we recommend the following for AIDS trials: (1) studies of long duration that allow evaluation of surrogate markers should receive high priority, (2) an ITT approach should be used for the analysis of a purely virologic end point whereby subjects who discontinue their randomly assigned treatment are followed up for virologic end points in the same manner as those continuing with the assigned treatment, and (3) in most trials, it may be important to analyze both a purely virologic end point and a regimen termination end point because they provide complementary information that rounds out the assessment of how the treatments should be used. The principle of analyzing both end points suggests that a purely virologic end point should be considered as the "default" primary end point of choice. An ITT analysis of a virologic end point guarantees that a secondary regimen termination analysis can be performed, while the converse is false. Regarding this point, for a virologic end point analyzed by ITT, sub-

jects who discontinue treatment are followed up until the occurrence of virologic events so that all regimen termination end point events will be captured. But if the primary end point is the regimen termination end point, subjects may not be followed up past treatment discontinuation, so some virologic end points will likely be missed.

Virologic Failure as an End Point

Many kinds of purely virologic primary end points have been used.[54] This end point is usually based on the time from randomization until plasma viral load level rises above a failure threshold (eg, 200 copies/mL) or by determining the proportion of subjects with adequate suppression up to or at a specified time point. For trials in which subjects enter with plasma HIV RNA levels above detection limits, a virologic end point is defined as the occurrence of either an early virologic failure, ie, a weak or absent virologic response or a rebound in viral load following a promising initial fall.

Limitations of Virologic End Points

The limited surrogacy of plasma HIV RNA levels was shown in a meta-analysis of all 16 randomized trials that compared outcomes involving nucleoside reverse transcriptase inhibitor regimens.[55] Trials with similar treatment-related 24-week changes in HIV RNA levels had widely varying treatment-related clinical outcomes.[55] It also failed to support the premise that HIV RNA markers reflect a treatment's effect on clinical outcomes to a larger extent than CD4 cell markers.[55]

Adult AIDS Clinical Trials Group (ACTG) trial 347[37] illustrates how we might be misled by using a purely virologic end point as though it were a clinical end point. Ninety-two subjects were randomized to receive amprenavir monotherapy or zidovudine-lamivudine-amprenavir triple therapy. Impetus for the trial was provided by prior studies.[56-61] The protocol-defined primary end point for efficacy was the proportion of subjects with virologic suppression to less than 500 copies/mL 24 weeks after randomization. An apparent high rate of virologic relapse early on triggered an interim review, which in turn led to closure of the amprenavir mono-

Table. Primary End Points Used in Adult AIDS Clinical Trials Group (ACTG) Antiretroviral Trials Opened to Accrual Between 1997 and 2000*

Study, y	ACTG Clinical Trial No.	Date Accrual Opened	Primary Objective of Trial	Virologic Suppression at Randomization	Primary End Point	Account for Early Virologic Failure	Analysis of Dropout†
Havlir et al,[36] 2000	343	February 1997	CVST maintenance	Yes (HIV RNA level <200 copies/mL)	Purely virologic	NA	DAC
Murphy et al,[37] 1999	347	March 1997	CVST new drugs	No	Purely virologic	Yes	DAC
Squires et al,[38] 2000	368	April 1997	CVST new drugs	No	Purely virologic	Yes	DAF
Gulick et al,[39] 2000	359	June 1997	CVST new drugs	No	Purely virologic	No	DAC
Albrecht et al,[40] 2000	364	July 1997	CVST new combinations	No	Purely virologic	No	DAC
Kuritzkes et al,[41] 2000	370	August 1997	CVST new combinations	No	Purely virologic	No	DAF
Adult ACTG Research Agenda Committees,[42] 2000	372A	September 1997	CVST maintenance/ intensification	Yes (HIV RNA level <500 copies/mL)	Regimen termination‡	NA	DAC
Hammer et al,[43] 1999	372B	September 1997	CVST new combinations	No	Purely virologic	Yes	DAF
25th ACTG Meeting Book,[44] 1998	388	June 1998	CVST intensification	No	Purely virologic	Yes	DAC
Smeaton et al,[45] in press	384	October 1998	Strategy	No	Regimen termination§	Yes	DAF
Hammer et al,[46] 2000	398	October 1998	CVST new combinations	No	Purely virologic	Yes	DAF
Adult ACTG Research Agenda Committees	A5025	November 1998	CVST maintenance/ intensification	Yes (HIV RNA level <200 copies/mL)	Regimen termination‡	NA	DAC
Adult ACTG Research Agenda Committees	400	December 1998	Strategy	No	Regimen termination§	Yes	DAF
25th ACTG Meeting Book	371	March 1999	CVST maintenance/ new drugs	Yes (HIV RNA level <200 copies/mL)	Purely virologic	NA	DAF
Adult ACTG Research Agenda Committees	A5064	November 1999	CVST intensification	No	Regimen termination§	Yes	DAF

*CVST indicates comparison of antiviral effect of specific treatments (ie, compares durability of virologic suppression between specific combination antiretroviral regimens); maintenance, the evaluation of a regimen's ability to maintain preexisting virologic suppression; NA, not applicable (in maintenance trials, the end point automatically does not include early virologic failure); new drugs, evaluation of new antiretroviral drugs; new combinations, comparison of regimens containing new combinations of drugs; intensification, comparison of regimens in which 1 or more of the regimens is designed to intensify the ability of a commonly used regimen to durably suppress viral replication; strategy, comparison of treatments using a predefined sequence of drug regimens; AIDS, acquired immunodeficiency syndrome; and HIV, human immunodeficiency virus.
†The method of handling dropout is either dropouts as censored (DAC), in which the event of withdrawal from the study is counted as censored, or dropouts as failures (DAF), in which the event of withdrawal is counted as failure.
‡Regimen termination end point 2 (defined in Box).
§Regimen termination end point 1 (defined in Box).

therapy arm. However, an ITT analysis of week 24 data provides the paradoxical impression that monotherapy was superior to the triple therapy arm, with 77% (30/39) vs 50% (20/40) of subjects having a plasma HIV RNA level of less than 500 copies/mL at 24 weeks, respectively (see Figure 3 of Murphy and colleagues[37]). The explanation is that all 39 subjects assigned to the monotherapy arm, many of whom failed virologically early in the follow-up period, were offered potent salvage regimens, compared with 32.5% (13/40) of those assigned to the triple therapy arm. Thus, ACTG trial 347 illustrates that a purely virologic end point can be a poor surrogate for clinical outcomes, especially when patients who fail with an inferior treatment subsequently receive a superior salvage treatment.

The need to switch regimens due to toxic effects reflects an important clinical effect of antiretroviral therapy that may be missed if a purely virologic end point is used. To illustrate this point, consider a typical trial in which 2 regimens are modestly effective, but a regimen of greater potency than either test regimen is offered as subsequent therapy for subjects having virologic failure or intolerable toxic effects with the initially assigned regimen. An ITT analysis could make it appear that a regimen with a high rate of severe toxic effects is superior because it leads to a quicker initiation of superior therapy (with resulting better virologic outcomes). In this situation, a regimen termination end point that counts treatment discontinuation as failure better addresses the trial's clinical objective than a purely virologic end point.

Drug resistance represents another important effect of antiretroviral therapy that is not fully measured by a purely virologic end point. For example, if in a 2-arm trial, the virologic failure rate is greater for regimen A than B but more subjects failing B have developed key resistance mutations, then it may be unclear which regimen is clinically preferable.[62] This example illustrates that for studies of single regimens, a purely virologic end point does not account for the resistance cost of having failed 1 treatment even if the patient is successfully suppressed with a second treatment. As elaborated in the next section, studies of sequences of therapies that use regimen termination end points for triggering treatment switches are needed for efficacy analyses to account for resistance costs.

Regimen Termination End Point

An ongoing trial that uses a regimen termination primary end point is ACTG trial 384[45] (Table, FIGURE). In this, an example of a strategy trial—defined as a comparison of approaches to using sequential regimens—subjects assigned to the strategy A arm receive efavirenz until virologic failure or treatment discontinuation and then receive nelfinavir. Subjects assigned to strategy B receive these regimens in the reverse order, and subjects assigned to strategy C receive 1 regimen including both efavirenz and nelfinavir (Figure). The primary end point is the time from randomization until both 3-drug regimens are terminated for strategies A and B or until the 4-drug regimen is terminated for strategy C. For this trial, the regimen termination end point is defined as the first occurrence of these events: virologic failure, permanent treatment discontinuation, AIDS-defining illness, death, and withdrawal from the study. This definition of the regimen termination end point is exhaustive, in that all outcomes other than successful virologic suppression with the assigned regimen are counted as failure events. This definition reflects the clinical question of which strategy keeps patients virologically suppressed and in the treatment program.

The ACTG trial 384 illustrates that a regimen termination primary end point may be appropriate in strategy trials that evaluate subjects for events beyond failure of initially randomized regimens through 2 or more sequential regimens. If strategies A and B show equal virologic suppression rates at the time of analysis but significantly more subjects assigned to the strategy A arm are on their second regimen, then strategy A may be inferior. This is true under the assumption that expending more treatment regimens by patients in strategy arm A places them at higher risk for clinical progression. In this case, analysis of a purely virologic end point would mislead by showing equality of the strategies, whereas analysis using a regimen termination end point would correctly show the inferiority of strategy A.

A competing variant of the regimen termination end point used by ACTG trial 384 is defined similarly except that subjects who withdraw from the study are censored at the time of withdrawal, for which censoring indicates that the time of failure is known only to exceed the time of withdrawal and the subject is considered to be successfully treated at the time. If dropout is unrelated to the risk of treatment failure then analysis of this end point is valid; however, if dropout is causally related to treatment failure, then analysis of the end point that considers dropout as failure gives an unbiased inference. Both methods likely miss the truth. For example, in a 2-arm trial, if dropout is associated with the ease of adherence and with an increased risk for virologic, toxic, or clinical events, then treating dropouts as failures biases the result toward the regimen that is easier to take, and censoring dropouts biases the result toward the regimen that is harder to take.[63] The goal of the trial (eg, intensification or simplification) determines which dropout mechanism makes the analysis conservatively biased toward the control arm. In most trials, it may be informative to use both approaches. The primary clinical question and knowledge of the regimens under study can guide the choice of primary end point. Many trials will benefit from sensitivity analyses of dropout assumptions.[64]

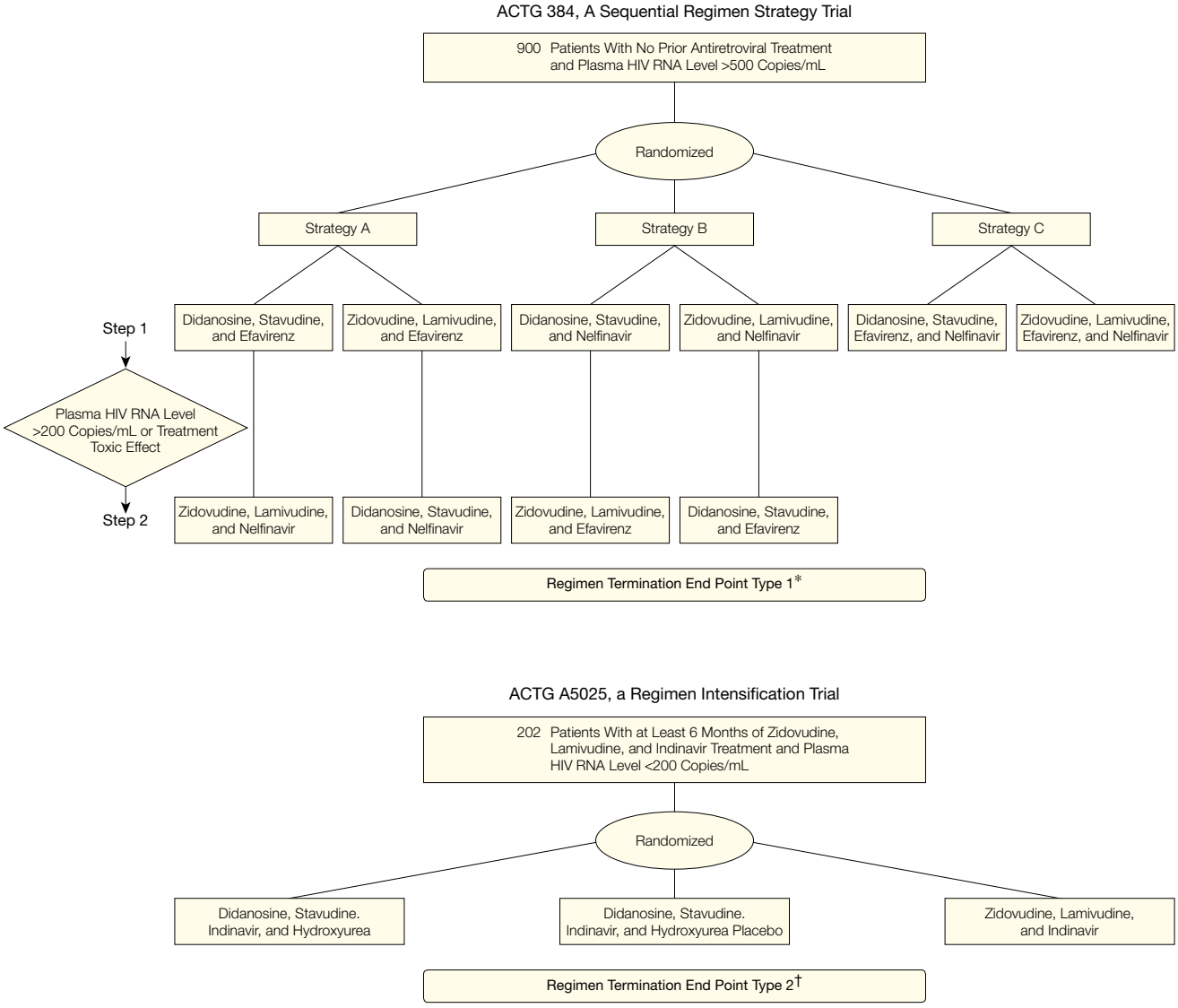
Another variant of the regimen termination end point includes as end points only virologic failure or treatment discontinuation due to confirmed protocol-defined toxic effects. This end point was used by ACTG trial A5025,[42] which tested the value of in-

tensifying a successful regimen (Table, Figure). All patients entering the trial were already receiving maximal benefit from existing therapy, so the tolerability of substitute regimens was essential in evaluating overall efficacy of a treatment switch. To reflect this goal, the regimen termination end point was selected for the primary analysis.

Comment

This article discusses 2 types of primary end points that have been used as surrogates for true clinical outcomes in antiretroviral trials. One end point is based purely on quantitative virologic information; the other is defined by fulfillment of the utility of a treatment regimen. We characterized the primary end point types that have been used within a particular clinical trials group by examining 15 Adult ACTG trials opened to accrual since 1997 (Table). We omitted industry trials because we were able to get a comprehensive sampling of the Adult ACTG trials. For interpretability, we thought it would be better to provide a complete list of the primary end

Figure. Examples of Trials Using Regimen Termination End Points: Adult AIDS Clinical Trials Group (ACTG) Trials 384[45] and A5025[42]



*The primary end point for ACTG trial 384 for strategy arms A and B was time from randomization until regimen termination after a switch to step 2 (triggered by a confirmed increase in plasma human immunodeficiency virus [HIV] RNA levels greater than 200 copies/mL or treatment discontinuation due to protocol-defined toxic effects); for strategy arm C, it was time from randomization until regimen termination. The composite regimen termination end point is defined by virologic failure, permanent treatment discontinuation, an acquired immunodeficiency syndrome (AIDS)–defining illness, death, or study withdrawal.
†The primary end point for ACTG trial A5025 was time from randomization until either a confirmed increase in plasma HIV RNA levels to greater than 200 copies/mL or discontinuation of randomized treatment due to protocol-defined toxicity.

points used within a particular clinical trials group rather than to also include an arbitrarily selected subset of primary end points that have been used in trials carried out by various companies. Eight of the 9 nonstrategy trials in which subjects had unsuppressed virus at the time of randomization used a purely virologic end point; whereas, 2 of the 4 trials in which subjects had suppressed virus at randomization used a regimen termination end point, and each of the 2 strategy trials used a regimen termination end point. This pattern reflects a preference for the purely virologic end point, except in settings such as strategy trials, in which the focus is more on different approaches to using sequential regimens than on comparison of specific drug regimens.

In addition to the paramount consideration of clinical relevance, when designing the primary end point for a trial, an investigator should consider the duration of the trial and the use of blinding. The duration of the trial should be long enough to observe enough primary end point events to reliably compare the treatments yet short enough so that the public receives the results in a timely manner. Thus, selecting candidate end points should be guided by expectations about event rates of the various end point types in the study population. For example, suppose a low rate of virologic failure events is expected and the studied regimens all provide durable virologic suppression so that the scientific focus is on tolerability. In this case, a regimen termination primary end point may be appropriate to protect against the trial continuing too long. Regarding blinding, the subjectivity of the end point evaluation increases with the amount of unblinding, which favors a purely virologic end point over the more subjective regimen termination end point.

For any given trial, the main criteria guiding the choice of surrogate end points include the primary study objective, the patient population, the objectivity of measuring the end points, and evidence (or clinical judgment) regarding the comparative accuracy of the end points as surrogates for true clinical outcomes. A purely virologic end point has the advantage of being able to be measured more objectively than a regimen termination end point, since patient-physician opinions about when to discontinue treatment determines the occurrence of the latter end point type but not of the former. When the goal of a study is to compare the virologic potency of specific drugs or regimens, a purely virologic end point is preferable; eg, in trials designed to assess the short-term activity of new drugs in early efficacy trials. In contrast, when the goal is to compare strategies for patient management (eg, sequencing of regimens), or when tolerability is considered to be essential to efficacy (eg, intensification of successful regimens), a regimen termination primary end point may merit consideration.

The approach to primary end point selection that we propose is based on hypotheses that have not been validated fully; validation would require showing that certain end points are better surrogates than others for clinical outcomes in certain settings. We acknowledge it would be interesting and important to provide an analysis assessing the association between a regimen termination end point and progression to clinical outcomes. However, to our knowledge, in all completed studies for which a regimen termination end point has been measured, the available follow-up data represent too brief a period to provide enough clinical events for a reasonably sensitive analysis. Long-term clinical outcome follow-up in randomized studies is needed for comparing the reliability of surrogate markers (the Adult ACTG is currently accruing a 5-year follow-up study involving thousands of subjects for this purpose). Data sets such as these will allow associations between surrogate end points and clinical end points to be studied as well as allow comparisons of the predictive surrogacy of various regimen termination and purely virologic end points. In such studies, it is important to compare several variants of purely virologic and regimen termination end points. It is also important to investigate the threshold for defining virologic failure, which must be prespecified for both end point types.[54] Using a low threshold (eg, 50 copies/mL) without evidence of its clinical relevance could lead to the discarding of useful therapies (D. V. Havlir, MD, unpublished data, 2000). A recent analysis of 2627 patients in the Swiss HIV Cohort Study[65] supports this concern, which showed comparable AIDS and death rates over 2.5 years in those maintaining suppression of less than 400 copies/mL vs those with viral rebound higher than 400 copies/mL following initial suppression. Lack of knowledge about meaningful thresholds is a major limitation of both end point types, especially for interim analyses because the decision rules for early termination depend on the selected threshold.

Prior to the completion of these crucial validation studies, the choice of primary end point necessarily is guided by current beliefs. If one believes that the effect of the investigated therapies on plasma HIV RNA levels captures the essential information needed to define the role of the therapies in clinical management for the target population, then a purely virologic end point is appropriate. Alternatively, if one believes that the need to switch regimens confers a higher risk of future disease progression than an increase in plasma HIV RNA level, then a composite regimen termination end point might be preferred. This belief supposes that the need to change regimens, and thus be exposed to the risk of multiple toxic effects and multidrug resistance, more closely measures tangible benefit (or lack thereof) for a patient than does an increase in viral burden alone. Until the data allowing for the definitive assessment of surrogacy are available, conducting analyses of both end points may be the wisest course since this may help in interpreting study results and applying them to clinical practice (eg, by providing an assessment of the relative amount of treatment-related difference in outcome due to virologic failure and to discontinuation due to treatment-related toxic effects).

For some study designs and patient populations, a biological marker other than a purely virologic one may be the best available surrogate end point. For example, in studies in which uniform virologic suppression is not the goal or is not realistically attainable, a CD4 cell outcome (or a combined HIV RNA and CD4 cell outcome) may be a viable primary end point candidate. Examples are treatment-interruption studies, in which HIV replication may fluctuate too much to be used as an end point, and salvage studies in those who have expended several regimens.

In cases in which the CD4 cell response is considered to provide predictive information for clinical outcomes beyond viral load alone (there is some evidence to support this),[10,66-68] again, the primary end point can include both biological markers. For example, failure of treatment can be defined as viral load above a threshold and CD4 cell count below a threshold, a situation in which the 2 thresholds may depend on each other. This end point is a purely biological marker end point, which does not need to be considered as a regimen termination end point; as mentioned, a regimen termination end point involves counting treatment discontinuation due to toxic effects as part of the treatment failure definition. The issue of how to handle both the biological markers of CD4 cell count and viral load jointly is independent of whether a regimen termination end point is used. The relevant point is that in some trial designs and populations, in particular, when discordant CD4 cell count and viral load responses are expected, an end point that includes both biological markers is a viable candidate for the primary end point.

When designing a trial, we recommend first considering a purely virologic end point as the primary end point and considering a regimen termination end point only if compelling arguments support it as a better surrogate for clinical outcomes or as better for addressing the practical clinical question. Using a composite primary end point without compelling reasons can unnecessarily complicate the separate evaluation of efficacy and safety, potentially complicating the regulatory process involved in drug approval, and can preclude the ability to carry out a well-powered ITT purely virologic secondary analysis.

The regimen termination end point can help clinicians balance efficacy and tolerability considerations when choosing a regimen for individual patients because it measures the average duration of the regimen's overall clinical utility. Assessing similarity or discordance between analyses of regimen termination and purely virologic end points helps clinicians weigh the efficacy/tolerability trade-offs. A regimen termination end point has incremental interpretative value over separate purely virologic end points and end points due to toxic effects. For example, analysis of the regimen termination end point helps distinguish whether an observed superiority of a regimen to suppress virus is due to increased tolerability or to having more drug options after discontinuation of the regimen.

We expect that regimen termination end points will be analyzed in most future trials, usually in secondary analyses in conjunction with primary analyses of purely biological marker end points and end points due to toxic effects, which will assist in interpreting study results. Used in isolation, regimen termination end points can be difficult to interpret and can potentially mislead because applying study results to populations is complicated by the end point's sensitivity to study clinician decision making and to the patient's ability to tolerate antiretroviral therapy. However, with the strategic use of drugs and drug combinations becoming an equally important issue as drug virologic activity, we envision that some studies will appropriately use a regimen termination primary end point.

In conclusion, selection of primary end points for AIDS trials is complicated by the long clinical course of the disease, the frequent onset of antiviral drug resistance, and limitations in data for validating surrogate end points. Five years of experience with potent antiretroviral therapies has suggested some concrete principles for end point selection (eg, clinical event rates have decreased in trials of antiretroviral therapy in which surrogate virologic primary end points were used). However, increasing the objectivity of the selection process in the future requires expansion of available information for the elucidation of the complex relationships between various surrogate end points and clinical end points. Only through vigilant collection of clinical outcomes data (eg, through routine collection of death event data from national death records) and data from long-term studies that monitor virologic, immunologic, and clinical information throughout sequences of regimens can this goal be achieved.

Funding/Support: This research was supported by grants 5-U01-AI38855, 5-U01-AI28076, and A1-32770 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

REFERENCES

1. Prentice RL. Surrogate endpoints in clinical trials. *Stat Med*. 1989;8:431-440.
2. Fleming TR, DeMets DL. Surrogate endpoints in clinical trials. *Ann Intern Med*. 1996;125:605-613.
3. Fleming TR. Evaluation of active control trials in AIDS. *J Acquir Immune Defic Syndr*. 1990;3:S82-S87.
4. Fleming TR. Evaluating therapeutic interventions. *Stat Sci*. 1992;7:428-456.
5. Lin DY, Fischl MA, Schoenfeld DA. Evaluating the role of CD4-lymphocyte counts as surrogate endpoints in HIV clinical trials. *Stat Med*. 1993;12:835-842.
6. Choi S, Lagakos SW, Schooley RT, Volberding PA. CD4+ lymphocytes are an incomplete surrogate marker for clinical progression in persons with asymptomatic HIV infection taking zidovudine. *Ann Intern Med*. 1993;118:674-680.
7. DeGruttola V, Wulfsohn M, Fischl M, Tsiatis A. Modeling the relationship between survival and CD4+ lymphocytes in patients with AIDS and AIDS-related complex. *J Acquir Immune Defic Syndr*. 1993;6:359-365.
8. Fleming TR. Surrogate markers in AIDS and cancer trials. *Stat Med*. 1994;13:1423-1435.
9. Hughes MD, Daniels MJ, Fischl MA, et al. CD4 cell count as a surrogate endpoint in HIV clinical trials. *AIDS*. 1998;12:1823-1832.
10. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann Intern Med*. 1997;126:946-954.
11. O'Brien TR, Blattner WA, Waters D, et al. Serum HIV-1 RNA levels and time to development of AIDS in the Multicenter Hemophilia Cohort Study. *JAMA*. 1996;276:105-110.
12. Mellors JW, Rinaldo CR Jr, Gupta P, et al. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science*. 1996;272:1167-1170.
13. Welles SL, Jackson JB, Yen-Lieberman B, et al. Prognostic value of plasma human immunodeficiency virus type 1 (HIV-1) RNA levels in patients with advanced HIV-1 disease and with little or no prior zidovudine therapy. *J Infect Dis*. 1996;174:696-703.

14. Katzenstein DA, Hammer SM, Hughes MD, et al. The relationship of virologic and immunologic markers to clinical outcomes after nucleoside therapy in HIV-infected adults with 200 to 500 CD4 cells per cubic millimeter. *N Engl J Med*. 1996;335:1091-1098.
15. Lee TH, Sheppard HW, Reis M, et al. Circulating HIV-1 infected cell burden from seroconversion to AIDS. *J Acquir Immune Defic Syndr*. 1994;7:381-388.
16. Coombs RW, Welles SL, Hooper C, et al. Association of plasma human immunodeficiency virus type 1 RNA level with risk of clinical progression in patients with advanced infection. *J Infect Dis*. 1996;174:704-712.
17. O'Brien WA, Hartigan PM, Martin D, et al. Changes in plasma HIV-1 RNA and CD4+ lymphocyte counts and the risk of progression to AIDS. *N Engl J Med*. 1996;334:426-431.
18. Phillips AN, Eron JJ, Bartlett JA, et al. HIV-1 RNA levels and the development of clinical disease. *AIDS*. 1996;10:859-865.
19. Hughes MD, Johnson VA, Hirsch MS, et al. Monitoring plasma HIV-1 RNA levels in addition to CD4+ lymphocyte count improves assessment of antiretroviral therapeutic response. *Ann Intern Med*. 1997;126:929-938.
20. Hammer SM, Squires KE, Hughes MD, et al. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. *N Engl J Med*. 1997;337:725-733.
21. Gulick R, Mellors J, Havlir D, et al. Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med*. 1997;337:734-739.
22. Collier AC, Coombs RW, Schoenfeld DA, et al. Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. *N Engl J Med*. 1996;334:1011-1017.
23. Mocroft A, Vella S, Benfield TL, et al. Changing patterns of mortality across Europe in patients infected with HIV-1. *Lancet*. 1998;352:1725-1730.
24. Palella FJ Jr, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N Engl J Med*. 1998;338:853-860.
25. Vittinghoff E, Scheer S, O'Malley P, et al. Combination antiretroviral therapy and recent declines in AIDS incidence and mortality. *J Infect Dis*. 1999;179:717-720.
26. Detels R, Munoz A, McFarlane G, et al. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. *JAMA*. 1998;280:1497-1503.
27. Hoyert DL, Kochanek KD, Murphy SL. Deaths: final data for 1997. *Natl Vital Stat Rep*. 1999;47:1-104.
28. Murray JS, Elashoff MR, Iacono-Connors LC, et al. The use of plasma HIV RNA as a study endpoint in efficacy trials of antiretroviral drugs. *AIDS*. 1999;13:797-804.
29. Staszewski S, DeMasi R, Hill AM, Dawson D. HIV-1 RNA, CD4 cell count and the risk of progression to AIDS and death during treatment with HIV-1 reverse transcriptase inhibitors. *AIDS*. 1998;12:1991-1997.
30. Wong JK, Hezareh M, Gunthard HF, et al. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science*. 1997;278:1291-1295.
31. Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. *MMWR Morb Mortal Wkly Rep*. 1998;47:43-82. Updates at: http://www.hivatis.org. Accessibility verified December 19, 2000.
32. Stone VE, Clarke J, Lovell J, et al. HIV/AIDS patients' perspective on adhering to regimens containing protease inhibitors. *J Gen Intern Med*. 1998;13:586-593.
33. Katzenstein DA. Adherence as a particular issue with protease inhibitors. *J Assoc Nurses AIDS Care*. 1997;8(suppl 8):10-17.
34. Yeh KC, Deutsch PJ, Haddix H, et al. Single-dose pharmacokinetics of indinavir and the effect of food [published correction appears in *Antimicrob Agents Chemother*. 1998;42:1308]. *Antimicrob Agents Chemother*. 1998;42:332-338.
35. Durant J, Clevenbergh P, Halfon P, et al. Drug-resistance genotyping in HIV-1 therapy. *Lancet*. 1999;353:2195-2199.
36. Havlir DV, Hellmann NS, Petropoulos CJ, et al. Drug susceptibility in HIV infection after viral rebound in patients receiving indinavir-containing regimens. *JAMA*. 2000;283:229-234.
37. Murphy RL, Gulick RM, De Gruttola V, et al. Treatment with amprenavir alone or amprenavir with zidovudine and lamivudine in adults with human immunodeficiency virus infection. *J Infect Dis*. 1999;179:808-816.
38. Squires K, Hammer S, De Gruttola V, et al. Randomized trial of abacavir in combination with indinavir and efavirenz in HIV-infected patients with nucleoside analog experience. From: 7th Conference on Retroviruses and Opportunistic Infections; January 30-February 2, 2000; San Francisco, Calif. Abstract 529.
39. Gulick RM, Hu XJ, Fiscus SA, et al. Saquinavir in combination with ritonavir or nelfinavir together with delavirdine, adefovir, or both in HIV-infected subjects with virologic failure on indinavir: ACTG 359. *J Infect Dis*. 2000;122:1375-1384.
40. Albrecht M, Katzenstein D, Bosch R, et al. ACTG 364-nelfinavir and/or efavirenz in combination with new NRTIs in nucleoside experienced subjects. From: 7th Conference on Retroviruses and Opportunistic Infections; January 30-February 2, 2000; San Francisco, Calif. Abstract 531.
41. Kuritzkes DR, Bassett RL, Johnson VA, et al. Continued lamivudine versus delavirdine in combination with indinavir and zidovudine or stavudine in lamivudine-experienced patients. *AIDS*. 2000;14:1553-1561.
42. The Adult AIDS Clinical Trials Group Research Agenda Committees. Available at: http://aactg.s-3.com/rac.htm#HIV. Accessibility verified December 29, 2000.
43. Hammer S, Squires K, DeGruttola V, et al. Randomized trial of abacavir (ABC) & nelfinavir (NFV) in combination with efavirenz (EFV) & adefovir dipivoxil (ADV) as salvage therapy in patients with virologic failure receiving indinavir (IDV). From: 6th Conference on Retroviruses and Opportunistic Infections; January 31-February 4, 1999; Chicago, Ill. Abstract 490.
44. *25th AIDS Clinical Trials Group Meeting Book*. Bethesda, Md: National Institute of Allergy and Infectious Diseases; August 1998.
45. Smeaton L, DeGruttola V. ACTG (AIDS Clinical Trials Group) 384: a strategy trial comparing consecutive treatments for HIV-1. *Control Clin Trials*. In press.
46. Hammer S, Mellors J, Vaida F, et al. A randomized, placebo-controlled trial of saquinavir, indinavir or nelfinavir in combination with amprenavir, abacavir, efavirenz, & adefovir in patients with protease inhibitor failure. From: 7th Conference on Retroviruses and Opportunistic Infections; January 30-February 2, 2000; San Francisco, Calif. Abstract LB7.
47. Dixon DO, McLaughlin P, Hagemeister FB, et al. Reporting outcomes in Hodgkin's disease and lymphoma. *J Clin Oncol*. 1987;5:1670-1672.
48. Anderson JR, Propert KJ, Harrington DP. Guidelines for reporting outcomes of lymphoma trials [letter]. *J Clin Oncol*. 1988;6:559-560.
49. Cardiac Arrhythmia Pilot Study (CAPS) Investigators. Effects of encainide, flecainide, imipramine and moricizine on ventricular arrhythmias during the year after acute myocardial infarction: the CAPS. *Am J Cardiol*. 1988;61:501-509.
50. Cardiac Arrhythmia Suppression Trial (CAST) Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med*. 1989;321:406-412.
51. International Chronic Granulomatous Disease Cooperative Study Group. A controlled trial of interferon gamma to prevent infection in chronic granulomatous disease. *N Engl J Med*. 1991;324:509-516.
52. Fisher LD, Dixon DO, Herson J, et al. Intention to treat in clinical trials. In: Peace KE, ed. *Statistical Issues in Drug Research and Development*. New York, NY: Marcel Dekker; 1990:331-350.
53. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials*. 3rd ed. New York, NY: Springer-Verlag; 1998.
54. Gilbert PB, Ribaudo HJ, Greenberg L, et al. Considerations in choosing a primary endpoint that measures durability of virologic suppression in an antiretroviral trial. *AIDS*. 2000;14:1-12.
55. HIV Surrogate Marker Collaborative Group. Human immunodeficiency virus type 1 RNA level and CD4 count as prognostic markers and surrogate end points: a meta-analysis. *AIDS Res Hum Retroviruses*. 2000;16:1123-1133.
56. Tisdale M, Myers RE, Maschera B, et al. Cross-resistance analysis of human immunodeficiency virus type 1 variants selected for resistance to five different protease inhibitors. *Antimicrob Agents Chemother*. 1995;39:1704-1710.
57. Partaledis JA, Yamagouchi K, Tisdale M, et al. In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to potent sulfonamide inhibitors of HIV-1 aspartyl protease. *J Virol*. 1995;69:5228-5235.
58. Sadler BH, Elkins M, Hanson C, et al. The safety and pharmacokinetics of 141W94: an HIV protease inhibitor. From: 5th European Conference on Clinical Aspects and Treatment of HIV Infection; September 27-29, 1995; Copenhagen, Denmark. Abstract 564.
59. Painter GR, Ching S, Reynolds D, et al. 141W94. *Drugs Future*. 1996;21:347-350.
60. St Clair MH, Millard J, Rooney J, et al. In vitro antiviral activity of 141W94 (VX-478) in combination with other agents. *Antivir Res*. 1996;29:53-56.
61. Schooley RT, the 141W94 International Study Group. Preliminary data on the safety and antiviral efficacy of the novel protease inhibitor 141W94 in HIV-infected patients with 150-400 CD4+ cells/mm 3. From: 36th Interscience Conference on Antimicrobial Agents and Chemotherapy; September 15-18, 1996; New Orleans, La. Abstract LB 7A.
62. Gilbert PB, Hanna J, DeGruttola V, et al. Comparative analysis of HIV type 1 genotypic resistance across antiretroviral trial treatment regimens. *AIDS Res Hum Retroviruses*. 2000;16:1325-1336.
63. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. New York, NY: Wiley; 1980.
64. Scharfstein DO, Rotnitzky A, Robins JM. Adjusting for nonignorable drop-out using semiparametric nonresponse models. *J Am Stat Assoc*. 1999;94:1096-1146.
65. Ledergerber B, Egger M, Opravil M. Clinical progression and virologic failure on highly active antiretroviral therapy in HIV-1 patients: a prospective cohort study. *Lancet*. 1999;353:863-868.
66. Grabar S, LeMoing V, Goujard C, et al. Clinical outcome of patients with HIV-1 infection according to immunologic and virologic response after 6 months of highly active antiretroviral therapy. *Ann Intern Med*. 2000;133:401-410.
67. Kaufmann D, Pantaleo G, Sudre P, Telenti A. CD4 cell count in HIV-1 infected individuals remaining viraemic with highly active antiretroviral therapy (HAART). *Lancet*. 1998;351:723-724.
68. Piketty C, Castiel P, Belec L, et al. Discrepant responses to triple combination antiretroviral therapy in advanced HIV disease. *AIDS*. 1998;12:745-750.