Guidance for Industry

Antiretroviral Drugs Using Plasma HIV RNA Measurements — Clinical Considerations for Accelerated and Traditional Approval

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

Clinical Medical

October 2002

Guidance for Industry

Antiretroviral Drugs Using Plasma HIV RNA Measurements — Clinical Considerations for Accelerated and Traditional Approval

Additional copies are available from:
Office of Drug Information (HFD-240),
Center for Drug Evaluation and Research,
Food and Drug Administration
5600 Fishers Lane,
Rockville, MD 20857
(Tel) 301-827-4573
http://www.fda.gov/cder/guidance/index.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

Clinical Medical

October 2002

TABLE OF CONTENTS

Guidance for Industry[1]

Antiretroviral Drugs Using Plasma HIV RNA Measurements — Clinical Considerations for Accelerated and Traditional Approval of

I. INTRODUCTION

This guidance is intended to assist sponsors in the clinical development of drugs for the treatment of human immunodeficiency virus (HIV) infection. Specifically, this guidance addresses the Agency's current thinking regarding designs of clinical trials that use HIV ribonucleic acid (RNA) measurements to support accelerated and traditional approvals of antiretroviral drug products. It is also intended to serve as a focus for continued discussions among the Division of Antiviral Drug Products (DAVDP), pharmaceutical sponsors, the academic community, and the public.

This guidance does not address specific phase-1 and phase-2 development issues, development of alternate dosing regimens, or the use of HIV-1 resistance testing. These issues will be addressed separately in future guidance documents.

In addition to consulting guidance documents, sponsors are encouraged to contact the division to discuss specific issues that arise during the development of an antiretroviral drug product.

II. BACKGROUND

Accelerated approvals of antiretroviral drugs have been based for years on changes in surrogate endpoints, such as CD_4 cell counts and plasma HIV RNA levels. Traditional approvals were based on clinical endpoint trials assessing the effects of a drug on mortality and/or HIV disease

[1] This guidance has been prepared by the Division of Antiviral Drug Products, Office of Drug Evaluation IV, in the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration.

progression. With the availability of potent antiretroviral drug regimens and sensitive assays for assessing plasma HIV RNA, the standards of clinical practice evolved to a paradigm emphasizing maximal and durable HIV RNA suppression. In addition, with the successes of combination therapy and the subsequent decline of HIV-related illnesses (Palella et al., 1998; Hogg et al., 1999), it became clear that a requirement for clinical endpoint studies for every traditional approval was no longer necessary nor feasible.

In July 1997, the Agency convened an advisory committee meeting to consider the use of changes in HIV RNA levels as endpoints in clinical trials supporting traditional approval of antiretrovirals. To evaluate the feasibility of using HIV RNA levels as a study endpoint, a collaborative group of pharmaceutical, academic, and government scientists investigated relationships between treatment-induced changes in HIV RNA and clinical endpoints from ongoing and completed antiretroviral trials (Murray et al., 1999; Hill et al., 1998). In several analyses of more than 5000 patients in multiple trials, a clear association was identified between initial decreases in plasma[2] HIV RNA levels and reduction in the risk of clinical progression and death. This relationship was observed across a range of patient characteristics including pretreatment CD4 counts and HIV RNA levels, prior drug experience, and treatment regimen. Based on these data, the Division of Antiviral Drug Products advisory committee concurred that treatment-induced decreases in HIV RNA *levels* were highly predictive of meaningful clinical benefit and that HIV RNA measurements could serve as endpoints in trials designed to support both accelerated and traditional approvals. The Division proposed that accelerated approvals could be based on studies that show a drug's contribution toward shorter-term reductions in HIV RNA (e.g., 24 weeks) while traditional approvals could be based on trials that show a drug's contribution toward durability of HIV RNA suppression (e.g., for at least 48 weeks). The committee agreed with this proposal and also recommended that changes in CD4 cell counts be consistent with observed HIV RNA changes when considering approval of an antiretroviral drug.

III. APPLICATIONS FOR ACCELERATED APPROVAL

A. Regulatory Definition of Accelerated Approval

According to the regulations (21 CFR 314.500 - 314.510), three criteria need to be addressed when considering the appropriateness of an accelerated approval: (1) the disease studied must be serious or life-threatening, (2) there must be an available surrogate that is reasonably likely to predict clinical benefit, (3) there must be demonstration of improved activity over approved drugs or activity in a population in need of additional therapeutic options. As stated in 21 CFR 314.500, accelerated approvals apply to drugs that "have been studied for their safety and efficacy in treating serious and or life-threatening illnesses and that provide meaningful therapeutic benefit to patients over existing treatments (e.g., ability to treat patients unresponsive to, or intolerant of available therapy, or improved patient response over available therapy)." An accelerated approval can be based on a surrogate endpoint reasonably likely to predict clinical benefit (but not necessarily a fully established surrogate) or a clinical endpoint

[2] Some trials measured HIV RNA in serum rather than plasma.

other than irreversible morbidity or mortality, where the ultimate goal of therapy is an effect on morbidity or mortality.

Under accelerated approval, continued marketing of a drug is subject to certain conditions outlined in the regulations, principally, an agreement to conduct further studies to establish clinical benefit. Since traditional approvals of antiretroviral drugs may now be based on changes in HIV RNA to establish benefit, shorter term reductions in HIV RNA levels supporting an accelerated approval can be considered *surrogate endpoints* for longer, durable suppression of HIV RNA levels.

Because continuous treatment with multiple antiretrovirals is necessary to achieve HIV suppression and because a substantial number of patients may no longer respond to available treatment options because of poor tolerability or the emergence of resistance, new antiretroviral agents are needed. Therefore, the Division strongly encourages sponsors to study patients who have few remaining approved treatment options. Demonstrating safety and drug activity (using a surrogate endpoint) in populations in need of new therapeutic options is consistent with the intended goals of the accelerated approval regulations**.**

For patients who have no remaining treatment options and are at imminent risk of disease progression secondary to severe immunodeficiency, participation in randomized controlled registrational trials may not always be desirable. For such patients, participation in open-label, compassionate-use safety studies or expanded access studies may be more appropriate. This was the consensus of the advisory committee panel held on January 2001. The Division encourages sponsors with promising new drugs to develop access protocols as soon as safely feasible in the drug development plan for patients in need. Trial design considerations for studies of heavily treatment experienced patients are addressed further in section III. C., Efficacy Considerations for Accelerated Approval.

The Division also recognizes that some investigational drugs may not be appropriate for patients with substantial amounts of prior antiretroviral drug exposure, particularly when an investigational drug has demonstrated broad cross-resistance to other drugs in its class. Therefore, other characteristics may also support a drug as a candidate for accelerated approval including improved efficacy or improved safety or tolerability over existing drugs. In addition, a more convenient dosing schedule (e.g., tid to bid or qd), novel mechanism of action, different clinical cross-resistance profile, favorable drug interaction profile, or utility in specific populations in need of therapies (e.g., pregnant women, children) may also support a drug as a candidate for accelerated approval. Such advantages should be demonstrable with data.

B. Safety Considerations for Accelerated Approval

The majority of antiretroviral accelerated approvals to date have been supported by safety data from at least 400 to 500 patients who received the proposed dose for marketing (or

higher doses) for approximately 6 months. Although HIV is considered life-threatening, the numbers of patients studied for previous antiretroviral accelerated approvals have approximated, or exceeded, the International Committee on Harmonisation (ICH) guidance for drugs intended for long-term treatment of *non-life-threatening* conditions.[3] The ICH guidance recommends the collection of safety data on at least 300 to 600 patients receiving the proposed dose for 6 months with safety data on a total of 1,500 patients when including patients with shorter-term drug exposures. The recommended safety database of 300-600 patients for 6 months was chosen to allow a reasonable chance to identify adverse events occurring at a frequency as low as 1:100. The ICH guidance also states that additional safety data on longer term use in a smaller cohort, than stated above, is advisable.

Applicants are encouraged to discuss their proposed safety database with the Division prior to submitting an NDA. On occasion, specific findings in preclinical or phase 1-2 development may indicate the need for a database that is larger or longer in duration to adequately evaluate potential drug toxicity.

Controlled and comparative safety data are preferred. Safety data from uncontrolled or expanded access protocols may be useful, but often lack the degree of detailed reporting obtained in controlled clinical trials. In addition, the assessment of causal relationships between a drug and an adverse event is more difficult when relying on uncontrolled safety data.

C. Efficacy Considerations for Accelerated Approval

1. General Design Issues

Every attempt should be made to design randomized, blinded (or partially blinded), controlled trials that permit one to clearly assess an investigational drug's contribution, as part of a combination regimen, toward decreases in HIV RNA. Studies in a broad range of patient populations (gender, age, and race) and a range of pretreatment characteristics (e.g., advanced and early disease, heavily pretreated and treatment naïve, as appropriate) are recommended to characterize the activity of the drug. The Division recommends that NDAs include at least two adequate and well-controlled studies of a minimum of 24 weeks duration to support accelerated approval. However, given that some patients will have longer term follow-up, submissions should include some data past 24 weeks when possible. In treatment naïve patients, analyses at earlier time points (e.g., 16 weeks) have proven to be less discriminating for detecting important differences between treatment regimens. In addition, prior to 24 weeks, some patients may have HIV RNA levels that are still declining, especially when measured with sensitive assays.

Blinded comparisons with controls are preferred because they reduce biases resulting from differences in management, treatment, or assessment of patients arising from investigator or subject knowledge of the randomized treatment. The Division

[3] ICH, *EIA The Extent of Population Exposure to Assess Clinical Safety: For Drugs Intended for Long-Term Treatment of Non-Life Threatening Conditions*, March 1995.

acknowledges that there are situations in which blinding drugs or regimens may not be feasible; however, in most cases the difficulties associated with blinding a study are not insurmountable. For example, blinding may be difficult when drugs require dose adjustments based on other components of a regimen; however, this could be accomplished by similarly dose adjusting the placebo. In combination regimens, the number of placebos could be large or unmanageable; however, in most cases blinding of one component of a regimen is all that is necessary. The Division encourages sponsors to provide access to their products' placebo formulations for use in clinical studies conducted outside their own drug development.

Sponsors designing studies in which blinding may be difficult or infeasible should discuss the proposal with the Division in advance, to review potential modifications that might facilitate blinding and to discuss the potential impact of open-label therapy on interpretation of results. When blinding is impossible, open-label protocols should have detailed procedures for treatment switches and toxicity management, since differential implementation of protocol procedures among treatment arms in open-label studies may impair interpretability of study results. For example, the validity of the results of open-label studies may be questioned if there are large differences between treatment arms with respect to nonprotocol-specified treatment discontinuations. In such instances DAVDP expects additional sensitivity analyses using different methods of handling treatment discontinuations or missing data, including those that treat the control and study arms asymmetrically (see Appendix B).

2. Study Population

Conducting controlled, comparative studies in patients who have exhausted many treatment regimens may call for innovations in study design, including the use of multiple investigational agents, factorial comparisons, and collaboration among two or more sponsors. Collaboration and the use of multiple investigational agents are strongly encouraged; however, phase-3 studies intended to support registration should be designed such that the treatment effect of each drug of interest can be isolated. In addition, the potential for drug-drug interactions, particularly those that may require dosing changes, should be considered in advance.

For heavily treatment experienced patients, the sponsor may consider studying doses that are higher than those studied in treatment naïve patients. Higher concentrations may be necessary to treat less susceptible HIV isolates. However, this approach may be limited by the amount of supporting preclinical data or by previously observed dose-limiting toxicities in the clinic.

As previously stated, patients with no approved treatment options and who are at imminent risk of progression may be better suited for expanded access or compassionate use protocols. Inclusion criteria for registrational studies should be such that patients are likely to complete the intended study period.

3. Study Design Options

- Superiority trial designs

Phase-3 superiority trials can include *add-on* or *substitution* comparisons. In the first case, the investigational drug plus a standard combination regimen is compared to placebo plus the same standard regimen. In some cases, an experimental drug could be added to a background regimen of drugs that the participant or investigator chooses from a list of possibilities.

For substitution comparisons, the investigational drug is substituted for a component of a standard regimen. This regimen is then compared to the standard regimen.

- Noninferiority trial designs

Noninferiority trials use substitution comparisons as described above. For noninferiority comparisons, it is important that the contribution of the substituted drug to a regimen's overall activity be previously characterized in the population of interest. This is often referred to as a study's *assay sensitivity*. This information should be used to support a noninferiority comparison and to calculate an appropriate sample size. This will be discussed further under section III. C. 3., Choice of Control Arms.

- Dose comparison trial designs

Phase 2 dose ranging studies that have demonstrated a significant dose response may provide supportive data for an accelerated approval of an antiretroviral drug. Generally, dose comparison studies should include a large enough range of doses to demonstrate a response slope. Given that it would not be desirable to design a large phase 3 protocol using doses of antiretrovirals that are known or predicted to be suboptimal based on preclinical or clinical data, phase 3 efficacy studies that primarily rely on dose comparisons may be difficult to design. Sponsors should discuss proposals for dose comparison studies with the Division in advance.

- Modified factorial comparison trial designs

This design is actually a variation of an add-on superiority design, except that multiple investigational agents may be tested simultaneously. This design may be useful when studying patients who are unable to construct a viable antiretroviral regimen from approved drugs.

A hypothetical example of a modified factorial is shown below. It is considered *modified* because not all possible cells of the factorial are studied. Excluding cells (e.g., background therapy alone) allows participants a greater possibility of being randomized to a treatment arm with more potentially active drugs. In the example below, the study tests superiority of arm 1 to arms 2 and 3 for two investigational drugs, X and Y. For drug Y, superiority of Arm 1 vs. 2 should be demonstrated. For Drug X, superiority of

Arm 1 vs. 3 should be demonstrated to support efficacy. This or alternative study designs using multiple investigational agents may be appropriate but should be discussed with the Division in advance.

Example A: Modified Factorial with Two Investigational Agents

Arm 1: Background Therapy + Investigational Drug X + Investigational Drug Y

Arm 2: Background Therapy + Investigational Drug X + Placebo for Drug Y

Arm 3: Background Therapy + Placebo for Drug X + Investigational Drug Y

It should be emphasized that background therapy may consist of other investigational agents available via expanded access or compassionate use mechanisms.

- Other Designs

As stated above, the Division acknowledges that designing adequate studies in heavily treatment experienced patients is difficult and may require modifications of typical study designs and procedures. Such issues were discussed during the January 2001 advisory committee meeting. In general, consensus was reached that use of external or historical controls would not be advisable because a large degree of heterogeneity in the treatment experienced patient population would impair the interpretability of such a comparison.

However, the committee did offer support for several similar proposals of a hybrid type design in which randomized comparisons were continued for short periods (days to weeks) followed by longer term assessments of the activity of a new drug as part of a combination regimen. There were several similar proposals. One example of such a hybrid design is shown below. Part 1 allows discrimination of the short-term antiviral effects of components of the regimen. Part 2 evaluates the durability of the antiviral effect produced by a combination regimen including investigational drug X. Further support for the investigational drug's (X) contribution to a durable antiviral effect could be provided by a correlation between baseline phenotypic susceptibility and antiviral response.

Example B: Two-Part Hybrid Design

Part 1 (Duration approximately 2 weeks)
Arm 1: Investigational drug X added to old regimen
Arm 2: Continue old regimen
Arm 3: Switch to new background regimen

Part 2: (22 weeks for a total of 24 weeks)
All Arms: Receive new background + investigational drug X

The duration of part 1 would depend on what is known regarding the rate of emergence of resistant strains under selective pressure with drug X and/or other drugs of the same class. Part 1 would allow a clean, albeit short, discrimination of the regimen's components toward the overall activity of the regimen.

A 24-week antiviral response rate that varied in accordance with baseline phenotypic susceptibility to drug X could offer supportive evidence that drug X was contributing to the regimen's overall efficacy at 24 weeks. For example, such a study might show that subgroups of patients with baseline phenotypic susceptibility to drug X of less than 5 fold of wild type, 5 to 10 fold of wild type, and greater than 10 fold of wild type had 24 week virologic response rates of 90 percent, 60 percent and 30 percent, respectively. This would be analogous to, though not as convincing as, a dose response. In this case patients are not randomized to dose but evaluated prospectively according to their baseline differences in susceptibility. However, this interpretation could be confounded if patients with better baseline phenotypic susceptibility to drug X also had better baseline susceptibility to other drugs in the regimen. In this case, the argument for using baseline susceptibility to support a particular drug's contribution toward activity would be weakened.

Because of the limitations of this design, such studies should not provide the principal support for an NDA package. However, such a study may complement a well-controlled study(ies) by providing important data for the use of the investigational drug in a population of need.

4. Choice of Control Arms

Every attempt should be made to include study treatment regimens consistent with standards of clinical practice while the trial is being conducted. In general, current HIV treatment guidelines emphasize the importance of using at least 3 potentially active drugs (if possible) when constructing a regimen. Proposals for control arms that deviate from current standards of care should be supported by convincing scientific rationale and/or data and discussed with the Division before implementation. Because of the evolving nature of accepted standards of HIV treatment, appropriate comparison regimens can be expected to change over time.

From a patient management perspective, use of control regimens that have been determined to be clearly suboptimal, as based on clinical studies or consensus of expert panels reviewing pertinent data, would jeopardize the viability of a study. From a regulatory perspective, controls used for efficacy comparisons should generally be those that have demonstrated durable viral suppression. Although one particular control regimen cannot be recommended as the most appropriate, some regimens are clearly inappropriate from either a regulatory or patient management perspective. For example, regimens consisting of only two of the currently approved nucleoside reverse transcriptase inhibitors (NRTI) in treatment naïve patients would not be considered appropriate control regimens from a patient management perspective.

Most antiretroviral drugs have not demonstrated activity in all situations or populations. In fact, some drugs are known to lack significant activity in some treatment experienced patients due to cross-resistance. However, for noninferiority trials, an *active* control must have reproducible and well-defined clinical activity in the context of the regimen and population to be studied. This is referred to as *assay sensitivity* for a clinical trial. In other words, the quantitative contribution of the control toward a regimen's overall activity should be previously demonstrated. Otherwise, a study drug may appear to be noninferior to an *active* control when in reality neither drug had contributed toward the activity of a regimen. Readers should refer to the guidance prepared by the ICH entitled, *E10 Choice of Control Group in Clinical Trials*. This document addresses key issues relating to the design of noninferiority studies.

5. Study Procedures

Protocols should include procedures for clinical management based on changes in HIV RNA. However, to facilitate interpretation of study results, it is critical that management decisions be made in a uniform manner. This is particularly important for open-label studies. Protocol procedures that allow treatment switches for patients who never achieve HIV RNA levels below an assay limit should be applied consistently across treatment arms. For example, some protocols allow treatment naïve patients who have not achieved an HIV RNA reduction of 1 $\log_{10}$ by 8 weeks to switch their antiviral regimen. These criteria may vary depending on the population studied and the response that is expected or desired.

6. Study Endpoints

It is anticipated that plasma HIV RNA measurements will be used as the primary assessment of drug efficacy. Sponsors are encouraged to use sensitive approved HIV RNA assays, particularly for studies in treatment naïve patients. More sensitive assays may allow greater discrimination of differences in treatment effects. In addition, the Division recommends the use of FDA approved HIV RNA assays to ensure that assay performance characteristics are understood. Currently, approved assays are the Roche Amplicor HIV-1 Monitor standard and ultrasensitive tests and the NucliSens HIV-1 QT. The Division prefers that approved assays be used. However, the Division also recognizes that new assays may offer potential improvements over existing technologies, perhaps with regard to sensitivity or detection of additional viral clades. In such cases, sponsors can choose to use unapproved assays but should be prepared to provide the Division with information supporting the limits and performance characteristics of the investigational assay (See Appendix A). This will permit an independent evaluation of an assay's limitations. Since the formal review of HIV RNA assays falls under the purview of the Center for Biologics, the Division's review of an unapproved assay's performance will focus on the interpretability of data with respect to the particular clinical trials in the NDA. Thus, the Division's review of assay performance data does not imply that the given assay is validated or FDA-approved for patient prognosis and/or monitoring. Furthermore, this review does not imply that the given assay is automatically acceptable for future clinical trials.

When feasible, the portion of subjects achieving HIV RNA levels below the assay limit at 24 weeks should be the primary endpoint for accelerated approval. Although such analyses are consistent with the current goals of clinical practice, it should be emphasized that the Division views this endpoint as a stringent tool for assessing drug activity and not necessarily a recommended threshold for patient management. The advantage of an endpoint that assesses response below an assay detection limit, compared to one assessing mean changes, is that it is more likely to be protected from treatment changes resulting from a perceived lack of efficacy. In addition, handling missing data is simplified when using a response endpoint compared to one assessing mean changes. However, the Division acknowledges that less stringent criteria or thresholds, other than confirmation of HIV RNA levels above an assay limit, may be useful for managing patients or for altering drug regimens.

Alternatively, for studies in patients who are unlikely to maximally suppress virus, responses other than those below an assay limit may be considered, such as a sustained 1 $\log_{10}$ decline in HIV RNA. Also, assessing mean changes in HIV RNA from baseline, including those that are averaged over time, may be useful for treatment experienced patients in which HIV RNA reductions below the assay limit are not expected to be frequent. Sustained HIV RNA reductions of at least 0.5 log have been shown to correlate with reductions in the risk of disease progression. Assessing mean changes from baseline are not considered as useful in patient populations that are able to achieve HIV RNA reductions below an assay limit. Sponsors are encouraged to discuss with the Division in advance which endpoint may be most appropriate for a particular protocol.

Analyses evaluating changes over time in CD_4 cell counts should accompany the analyses of HIV RNA. Clinical endpoint data (CDC class C events) should also be collected, analyzed, and submitted with the NDA. However, the frequency of such events is likely to be low particularly over a 24-week time period.

7. Statistical Considerations

The choice of delta for noninferiority testing should be discussed with the Division prior to study initiation because one delta is not appropriate for all study designs. In the past, many noninferiority studies have been powered based on a delta of 10 percent to 12 percent. In most cases, this allowed sponsors to plan studies that would be reasonably powered and feasible to conduct. Such a delta has been useful when comparing the most potent component of a three-drug regimen in treatment naïve individuals. However, the sponsor should ultimately attempt to choose a delta based on prior knowledge of the quantitative contribution of the active control (substituted part of the drug regimen) to the regimen as a whole. This contribution should be determined in a similar population with a similar length of follow-up of the proposed study. For noninferiority testing, sponsors should employ two-sided 95 percent confidence intervals adjusted for multiple comparisons. If one-sided confidence intervals are used, the alpha should be 0.025.

Both noninferiority and superiority can be assessed in the same study provided that the noninferiority comparison and choice of delta have been specified prior to study initiation and/or provided that the choice of delta can be clearly justified based on previous clinical data.

Analyses that include all randomized patients should be included in all NDAs. Sometimes patients who were randomized but never returned to receive drug, can be removed from the analyses if the study is blinded or if both arms are similar in this regard. In such analyses, patients who have introduced a new antiretroviral drug to the regimen (except for prespecified changes in the background regimen in which the reason for change is due to the background drug and not due to the study drug or control) have discontinued study, have been lost to follow-up, or for whatever reason have missing HIV RNA data should be considered to have HIV RNA levels above the assay limit. This is sometimes referred to as a *noncompleter equals failure (NC=F)* analysis. In general, missing HIV RNA data between study visits with values below the assay limit does not constitute treatment failure.

In addition to a NC=F analysis, an analysis comparing documented virologic failures in evaluable patients should also be submitted. Any inconsistencies in treatment outcomes for these analyses should be explained. Sensitivity analyses that use different methods of handling treatment discontinuations and missing data should also be provided in support of efficacy.

8. Reporting Efficacy Results

The Division prefers that study results be reported in a manner in which the reason for treatment failure can be easily discerned. For accelerated approvals, simple proportions may be calculated for the possible treatment outcomes at 24 weeks as shown in the example in Table 1. Refer to Appendix B for computing treatment outcomes and displaying 48-week data for traditional approval.

Table 1. Example of Data Display of Treatment Outcomes at 24 weeks (accelerated approval)

Outcome at 24 weeks	**Experimental N (%)**	**Control N (%)**
Below assay limit	xx (xx%)	xx (xx%)
Above assay limit	xx (xx%)	xx (xx%)
Death	xx (xx%)	xx (xx%)
Drug change or discontinuation due to adverse events	xx (xx%)	xx (xx%)
Drug change or discontinuation due to other reasons	xx (xx%)	xx (xx%)
Consent withdrawn	xx (xx%)	xx (xx%)
Loss to follow up	xx (xx%)	xx (xx%)

Nonadherence	xx (xx%)	xx (xx%)
Pregnancy	xx (xx%)	xx (xx%)
Protocol violation	xx (xx%)	xx (xx%)
Other	xx (xx%)	xx (xx%)

IV. APPLICATIONS FOR TRADITIONAL APPROVAL

In this guidance, the term *traditional approval* refers to the usual marketing clearance mechanism for the majority of drugs that have demonstrated clinical efficacy as shown by a treatment effect on a clinically meaningful endpoint. In most cases, results from at least two well-controlled studies should be included in an NDA to support traditional approval. The guidance entitled *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* (May 14, 1998) should be consulted for discussion of circumstances in which approval may be considered on the basis of a single trial.

A. Safety Considerations for Traditional Approval

It is expected that traditional approvals will be supported by safety data from a minimum of 500 patients who have received a drug for approximately 48 weeks. In addition, sponsors should provide any safety data from earlier study enrollees who have been followed for periods longer than 48 weeks.

Because multiple adverse events have been observed with chronic administration of antiretroviral therapy, mechanisms should be used for systematically evaluating adverse events over prolonged periods following traditional approval. Controlled comparisons and prospectively evaluated cohorts may be helpful in characterizing and defining drug associations for late-occurring adverse events. Therefore, after traditional approval, the Division strongly encourages sponsors to continue to collect safety data in key randomized studies or other treatment cohorts for prolonged periods (3-5 years). For long-term follow-up, study procedures and data collection may be targeted or streamlined to relieve some of the burden of long-term safety assessments.

Controlled and comparative safety data are preferred. Safety data from uncontrolled compassionate use protocols may be supportive, but often lack the degree of detailed reporting obtained in controlled clinical trials. Uncontrolled safety data rarely allow one to assess causal relationships between a drug and an adverse event.

Since a risk-benefit ratio is always considered in any approval, careful attention should be paid to treatment discontinuations for intolerance or toxicity. Sponsors should make every attempt to ascertain reasons for treatment and/or study discontinuations. In addition, patients who did not have a protocol-defined dose-limiting toxicity, but nonetheless had an unresolved intolerance or adverse event at the time of discontinuation, should be classified as discontinuing treatment secondary to drug intolerance and not due to *patient's choice* or *other*. Analyses should be performed to evaluate reasons for

treatment discontinuations, possible baseline risks for treatment intolerance, and time until a dose-limiting adverse reaction. Such analyses are particularly crucial in studies with a substantial proportion of treatment discontinuations (greater than 20-25 percent), or with different rates of discontinuations among treatment arms.

B. Efficacy Considerations for Traditional Approval

1. Study Design

The same studies that are evaluated at 24 weeks for accelerated approval may be continued for 48 weeks and longer to support traditional approval. The duration of these studies should permit the last patient randomized and still continuing therapy to have passed the 48-week time point. The final study report should include all available data at the time of analysis, including that beyond 48 weeks. Some studies should be continued beyond traditional approval, as feasible, as part of a post-marketing commitment to assess safety and efficacy of chronic administration. Data regarding longer term virologic assessments in principal studies can be displayed in drug labeling.

Issues relating to choice of control arms, comparisons, and study procedures are discussed in previous sections.

2. Study Endpoints

Traditional approvals can be based on study results that show the drug's contribution toward sustained suppression of plasma HIV RNA levels. However, some applications can contain a combination of clinical endpoint studies and HIV RNA studies. Thus, the types of studies included will affect the wording of the indications granted.

The sponsor should also include supportive analyses of CD_4 count responses and clinical endpoints. In virologic studies, the investigational drug should also show no deleterious effect on clinical progression (CDC-C AIDS-defining illnesses and death); however because the study is unlikely to have sufficient power to detect treatment differences in clinical events, statistically significant differences between treatment arms are not expected. In addition, studies generally should show favorable CD_4 responses; however, there may be clinical situations in which significant CD_4 increases may not be demonstrated, such as in patients who already have normal range CD_4 counts.

Study results showing discordant CD_4 responses will require close examination and if not explained could call the primary endpoint into question. Drugs that produce major discordance between the desired HIV RNA and CD_4 cell count responses (i.e., a drug that produces decreases in both HIV RNA and CD_4 cell counts instead of decreases in HIV RNA with increases in CD_4 cell counts) should probably be evaluated using clinical endpoint studies.

The proportion of patients with HIV RNA levels below the assay limit at 48 weeks (or longer) and time-to-loss-of-virologic-response will be considered primary endpoints for

trials supporting traditional approval. These virologic response endpoints are discussed below.

- Time-to-loss-of-virologic-response

Studies designed to assess time-to-loss-of-virologic-response allow participants who have lost a virologic response to switch or seek new therapy without compromising major study analyses. One definition for time-to-loss-of-virologic-response is the time between randomization (or start of treatment) and the last value below an assay limit in a patient who initially suppressed HIV RNA below an assay limit but subsequently demonstrated virologic rebound (two consecutive HIV RNA levels above the assay limit). Subjects who do not achieve suppression below the assay limit during the study (or within a predefined shorter time period allowing for earlier treatment switching) can then be defined as having a time to loss of response of zero. See appendix A for an algorithm that the Division typically uses to define loss of virologic response. Other algorithms may be appropriate depending on the study designs. Analysis of the total duration below the assay limit may also be presented, usually as a secondary analysis.

For the primary analysis, the Division usually considers patients who die, are lost to follow-up, or introduce new study treatment due to toxicity/intolerance (or any other reason) as treatment failures at the time of those events. Patients who have introduced a new antiretroviral drug regimen are also usually considered to be nonresponders unless such changes only involve alterations in background therapy that are permitted in the protocol. Other analyses based on virologic failure alone should also be performed. Patients who experience a CDC-C clinical event, but who otherwise are maintaining complete viral suppression on randomized therapy, can be considered to be responding.

- Proportion with HIV RNA below an assay limit of quantitation

Since statistical methods are insufficient for time to event analyses for noninferiority comparisons, assessing proportions of patients below the assay limit at a prespecified time point is recommended for noninferiority trials. However, the Division prefers the use of Kaplan Meier plots (using the algorithm in appendix B) to estimate the proportions of treatment responders at 48 weeks. See Statistical Considerations below.

- Clinical endpoints

Adequate and well-controlled trials showing clinical benefit, as measured by the occurrence of new AIDS defining events (CDC Class C events) or death, will continue to be considered acceptable support for traditional approval. Results of such studies can be described in the package insert, and could influence the approved indication(s).

It is imperative that all new Class C CDC-defined events be thoroughly documented and analyzed in registrational trials, even those in which the primary endpoint is virologic. Since clinical events may sometimes be difficult to assess, the Division recommends that all study protocols have independent assessment and adjudication of all Class C events,

using case report forms and additional medical records to fully document their occurrence when appropriate.

C. Statistical Considerations

The same considerations for trials supporting accelerated approval also apply to traditional approval. It should be emphasized that studies for traditional approval should be analyzed after the last patient randomized has completed 48 weeks of treatment (if still on therapy). Therefore, for many participants, there will be data points past 48 weeks. As much extended data as possible should be included and evaluated in the NDA.

When assessing superiority for time to loss of virologic response, the log rank test for differences in the Kaplan-Meier (KM) curves, using all available follow-up data, should be performed. For noninferiority comparisons, differences in the KM proportions below the assay limit at a prespecified time point with associated confidence intervals should be assessed.

REFERENCES

Chuang-Stein, C., and R. Demasi, 1998, "Surrogate Endpoints in AIDS Drug Development: Current Status," *Drug Information Journal,* 32:439-448.

Hill, A.M., R. DeMasi, and D. Dawson, 1999, "Meta-Analysis of Atiretroviral Effects on HIV-1 RNA, CD_4 Cell Count and Progression to AIDS or Death," *Antiviral Therapy,* 3:139-145.

Hogg, R.S., B. Yip, and C. Kully, 1999, "Improved Survival Among HIV-Infected Patients After Initiation of Triple-Drug Antiretroviral Regimens," *CMAJ,*160:659-65.

Marschner, I.C., A.C. Collier, R.W. Coombs, et. al., 1998, "Use of Changes in Plasma Levels of Human Immunodeficiency Virus Type 1 RNA to Assess the Clinical Benefit of Antiretroviral Therapy," *J Infect Dis,* 177:40-7.

Murray, J.S., M.R. Elashoff, L.C. Iacono-Connors, et al., 1999, "The Use of Plasma HIV RNA as a Study Endpoint in Efficacy Trials of Antiretroviral Drugs," *AIDS,*13:797-804.

Palella, F.J., K.M. Delaney, A.C. Moorman, et al., 1998, "Declining Morbidity and Mortality Among Patients with Advanced Human Immunodeficiency Virus Infection, *N. Engl. J. Med*., 338:853-60.

APPENDIX A: REGULATORY CONSIDERATIONS WHEN USING NEW HIV RNA ASSAYS

A. General Considerations

In this guidance a *new* assay refers to any assay that has not been approved by the FDA or to an approved assay that is being used in a manner different than described in its labeling. The assay that is used in the clinical trial should be identical to the assay that is used to assess the performance characteristics. Clinical studies can use HIV RNA measurements either to quantify the amount of HIV RNA in patient samples (e.g., copies/mL) or to classify a patient sample as above or below a specific value. Therefore, considerations for both quantitative and qualitative uses will be addressed in the following subsections.

It is recommended, but not required, that FDA-approved HIV RNA assays be used to support clinical trials. Currently approved assays are the Roche Amplicor HIV-1 Monitor standard and ultrasensitive tests and the NucliSens HIV-1 QT. However, when experimental and/or investigational HIV RNA measurement assays are used to support clinical trials, sufficient data characterizing assay performance should be provided. This permits an independent evaluation of an assay's limitations. Review of assay performance by the Division will focus on the interpretability of data generated by the assay with respect to the particular clinical trials in the NDA. Thus, the Division's review of assay performance data does not imply that the given assay is validated, or FDA-approved, for patient prognosis and/or monitoring. Furthermore, this review does not imply that the given assay is automatically acceptable for future clinical trials.

B. Investigational HIV RNA Assays

Assay design rationale, essential methodology, and performance characteristics are important components of information submitted to support new assays. Assay performance characteristics studies should be conducted on specimens that are representative of the HIV target subtype (Clade/s), and from the same tissue reservoirs (serum, plasma, other) assessed in the clinical trial. Specimen stability (handling, processing, and storage protocols) data should show no significant change of HIV RNA material as measured by the assay. Generally, these data should be derived from protocol-based experiments. Protocols and quality assurance/quality control information for the assay should be submitted with the data.

1. Quantitative Assays: Performance Studies

To support clinical virology data that rely on quantitative assessments, the methodology and/or technology used to generate those data should be adequately described in the application. An HIV RNA quantitative assay should be able to accurately and precisely report HIV RNA copy numbers over a defined range. Assay performance characteristics should include, but not be limited to, information and/or data that define the assay accuracy, precision, sample stability, and effects of certain interfering substances.

Accuracy can be assessed by calculating the mean of repeated observations of a given known sample and comparing the mean to the known input value. Assay accuracy

should be determined across the proposed range of the assay. Precision can be assessed by calculating the mean square error (MSE) and converting to a percent coefficient of variation (CV). Precision should be determined across the proposed range of the assay. The quantitative limit of the assay will be determined by the lowest input value where the assay maintains its accuracy and precision. A quantitative upper limit may be similarly defined.

Ultimately, the quantitative limit should be supported by data that characterized the assay performance characteristics. Laboratory strains and unique clinical HIV test specimens should be used to derive the performance characteristics of the assay. Each of these test specimens should first be independently and adequately quantitated (i.e., by comparability to an acceptable standard) prior to being used to define the performance characteristics of the new assay.

2. Qualitative Assays: Performance Studies

Assays that are used in a qualitative manner should have the ability to distinguish between known HIV seropositive clinical specimens and known HIV seronegative specimens with 95 percent confidence. A threshold or screening cut-off value (qualitative limit), expressed in HIV RNA copy numbers per mL should be determined. An assay result would be expressed as either a ≥ or < the HIV RNA copy number qualitative limit. However, a result that is below the screening cut-off value does not imply that the specimen is HIV negative, it implies only that the specimen has less virus material than that needed to distinguish the specimen from a known negative with 95 percent confidence. Assay performance characteristics should include, but may not be not limited to, information and/or data that define the assay range of specificity, range of sensitivity, sample stability, and effects of certain interfering substances.

The Division recommends that the range of specificity of the assay be defined as the 95 percent confidence interval of reported observations from 500 random seronegative blood or plasma donors. The range of sensitivity of the assay may be defined as the 95 percent confidence interval of reported observations from 200 unique seropositive samples. Each of these seropositive samples should be quantified by an independent method and then diluted to the proposed qualitative limit prior to assessing the assay range of sensitivity. It is expected that the two ranges will not overlap.

Ultimately, the qualitative limit can be supported by and derived from the assay performance data. Laboratory strains and unique clinical HIV test specimens should be used to derive the performance characteristics of the assay. Each of these test specimens should first be independently and adequately quantitated (i.e., by comparability to an acceptable standard) prior to being used to define the performance characteristics of the new assay.

C. Glossary of Assay Terminology

Quantitative Assay: An assay that is accurate and precise over a defined range.

Qualitative Assay: An assay that can distinguish between a known HIV positive specimen and a HIV seronegative specimen.

Range of Specificity: A 95 percent confidence interval of reported observations from 500 seronegative random blood or plasma donors.

Range of Sensitivity: A 95 percent confidence interval of reported observations from dilution of 200 unique seropositive samples to the proposed qualitative limit, each quantified prior to dilution by an independent method.

Quantitative Limit: The lower boundary of the accurate and precise defined range.

Qualitative Limit: The lowest concentration of HIV RNA copies per mL that the assay can reliably distinguish from seronegative samples.

Interfering Substances: Any substance and/or infectious agent that may be present in a clinical sample and affect a performance characteristic of the new assay.

Precision: The variability in terms of the mean square error (MSE) converted to a percent CV (CV= the square root of MSE divided by the expected value x 100 percent) within the proposed range.

Accuracy (Bias): The mean of repeated observations of a given known sample compared to the expected value for knowns within the proposed range of the assay.

Sample stability: Specimen handling, processing, and storage procedures that result in no significant changes in expected HIV RNA copy numbers.

APPENDIX B: TIME TO LOSS OF VIROLOGIC RESPONSE ALGORITHM

- For NDAs with 48-week virologic data, analyses comparing time to virologic failure can be assessed using the following algorithm:

1. For 2 and 3 below, discard all visits with no data. In what follows, *visit* means visit with an observed viral load. All available visits, including off-schedule visits and post-week 48 visits, should be used for the calculation. Data should not be interpolated for visits or time points with missing data.
2. Subjects who never achieved confirmed HIV RNA levels below the assay limit (on two consecutive visits) before any of the following events will be considered to have failed at time 0.
    a. Death
    b. Introduction of a new antiretroviral drug to the regimen. Exceptions may be made for certain prespecified changes in background therapy where the reason for change is due to toxicity or intolerance of background therapy and not the study drug or control.
    c. Last available visit
3. For all subjects who have confirmed HIV RNA levels below an assay limit (two consecutive visits below an assay limit), the time to failure is the earliest of the choices below, with modification specified in 4:
    a. Time of the event as described in 2b
    b. Time of loss to follow-up
    c. Time of confirmed levels above an assay limit. Confirmed is defined as two consecutive visits greater than an assay limit or one visit greater than an assay limit followed by loss to follow-up.
    d. Time of death
4. If the time to virologic failure defined above is immediately preceded by a single missing scheduled visit or multiple consecutive missing scheduled visits, the time of virologic failure is replaced by the time of the first such missing visit.

For open-label studies, algorithms that incorporate other ways of handling missing data or treatment changes may be used for additional sensitivity analyses. For example, sponsors should perform analyses that treat non-protocol-specified treatment changes as failures in the study arm and as censored at the time of change in the control arm when exploring the sensitivity of the results to potential biases related to an open-label design.

- Computation of proportions of success and display of data

The reasons for failure at week 48 should be the primary reason for the earliest treatment failure. For any visit, subjects who have achieved confirmed virologic success (two consecutive visits below and assay limit) but have not failed according to loss of virologic response according to the above algorithm should be considered successes. All others should be classified into the

other appropriate categories listed in Table 2. Visit windows should be prespecified in the protocol, and all events occurring in the window should be considered events for that visit.

Display of 48-week outcome as calculated using the above algorithm

Outcome at 48 weeks	**Experimental N (%)**	**Control N (%)**
Below assay limit	xx (xx%)	xx (xx%)
Above assay limit (confirmed) **Rebound** **Never suppressed through week 48** **Drug change due to virologic failure**	xx (xx%)	xx (xx%)
Death	xx (xx%)	xx (xx%)
Drug change or discontinuation due to adverse events	xx (xx%)	xx (xx%)
Drug change or discontinuation due to other reasons	xx (xx%)	xx (xx%)
Consent withdrawn	xx (xx%)	xx (xx%)
Loss to follow up	xx (xx%)	xx (xx%)
Nonadherence	xx (xx%)	xx (xx%)
Pregnancy	xx (xx%)	xx (xx%)
Protocol violation	xx (xx%)	xx (xx%)
Other	xx (xx%)	xx (xx%)

*Subjects who were never suppressed but were loss to follow-up or introduced new drug therapy before week 48 should be classified into drug change or discontinuation categories. Prespecified background drug changes where the reason for change is due to the background drug can be ignored in the analysis with consent from the Division.