Guidance for Industry

Human Immunodeficiency Virus-1 Infection: Developing Antiretroviral Drugs for Treatment

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact Jeffrey Murray at 301-796-1500.

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

June 2013
Clinical Antimicrobial

Revision 1

Guidance for Industry

Human Immunodeficiency Virus-1 Infection: Developing Antiretroviral Drugs for Treatment

Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 51, rm. 2201
Silver Spring, MD 20993-0002
Tel: 301-796-3400; Fax: 301-847-8714; E-mail: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

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

June 2013
Clinical Antimicrobial

Revision 1

TABLE OF CONTENTS

Guidance for Industry[1]
Human Immunodeficiency Virus-1 Infection: Developing Antiretroviral Drugs for Treatment

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. INTRODUCTION

This guidance provides recommendations for the development of antiretroviral drugs regulated within the Center for Drug Evaluation and Research at the Food and Drug Administration (FDA) for the treatment of human immunodeficiency virus-1 (HIV-1 or HIV) infection.[2] Specifically, this guidance addresses the FDA's current thinking regarding the overall development program and clinical trial designs for antiretroviral drugs to support an indication for the treatment of HIV-1 infection. This draft guidance is intended to serve as a focus for continued discussions among the Division of Antiviral Products (DAVP), pharmaceutical sponsors, the academic community, and the public.[3] The organization of the guidance parallels the development plan for a particular drug or biologic.

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

[2] For the purposes of this guidance, all references to *drugs* include both human drugs and therapeutic biological products unless otherwise specified.

[3] In addition to consulting guidances, sponsors are encouraged to contact the division to discuss specific issues that arise during the development of antiretroviral drugs.

This guidance revises the guidance for industry *Antiretroviral Drugs Using Plasma HIV-RNA Measurements — Clinical Considerations for Accelerated and Traditional Approval* issued in October 2002.[4] After it has been finalized, this guidance will replace the October 2002 guidance. Significant changes from the 2002 version include: (1) more details on nonclinical development of antiretroviral drugs; (2) a greater emphasis on recommended trial designs for HIV-1-infected heavily treatment-experienced patients (those with multiple-drug resistant virus and few remaining therapeutic options); (3) use of a primary endpoint evaluating early virologic changes for studies in heavily treatment-experienced patients; and (4) use of the traditional approval pathway for initial approval of all antiretrovirals with primary analysis time points dependent on the indication sought instead of an accelerated approval pathway followed by traditional approval.

This guidance does not address the use of antiviral drugs for preventing the transmission of HIV-1 infection. Also, this guidance does not address the development of therapeutics, without antiviral mechanisms, intended to mitigate or reverse clinical or pathophysiological outcomes of immunologic suppression of HIV-1 infection.

Additionally, this guidance does not contain discussion of the general issues of clinical trial design or statistical analyses for HIV antiretroviral trials. Those topics are addressed in the ICH guidances for industry *E9 Statistical Principles for Clinical Trials* and *E10 Choice of Control Group and Related Issues in Clinical Trials*. This guidance also does not contain details regarding nonclinical safety and toxicology studies that should be conducted in standard animal models as described in the guidance for industry *Nonclinical Safety Evaluation of Drug or Biologic Combinations.*

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

Brief summaries of HIV infection and treatment and the regulatory history of antiretroviral drug development and approvals are included below to support the rationale for changes in antiretroviral drug development guidance.

HIV Infection and Treatment

HIV infection is a chronic viral infection that, when untreated, causes a progressive destruction of the immune system resulting in acquired immunodeficiency syndrome (AIDS). The key component of the immune deficiency associated with untreated HIV

[4] We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA Drugs guidance Web page at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

replication is a marked reduction in cluster of differentiation 4 (CD4) T-cells, but derangements in other immunologic parameters also play a role in the immune deficiency syndrome. AIDS is defined as the presence of HIV infection with a CD4 cell count less than 200 cells/mm^3 and/or the presence of an AIDS-defining clinical condition, which includes any number of opportunistic infections, malignancies, or other clinical syndromes as defined by the Centers for Disease Control and Prevention (CDC 1992).

Current treatment of HIV consists of a combination of antiretroviral drugs referred to as Highly Active Antiretroviral Therapy (HAART). HAART typically consists of three antiretroviral drugs from two or more drug classes. Sometimes more than three drugs are used in patients who have been treated previously and are known or presumed to harbor viral strains with reduced susceptibility. In addition, some HAART regimens include a drug that increases or prolongs exposures of one or more drugs in the regimen because of an intentional drug interaction. Such a drug is referred to as a pharmacokinetic (PK) *booster* or a PK *enhancer*.

The goal of antiretroviral treatment is to indefinitely maintain suppression of plasma HIV-RNA levels (also called viral load) below the detection limits of sensitive HIV-RNA assays. For initiating first-line therapy in treatment-naïve patients, several guidelines recommend preferred regimens. Current preferred regimens in treatment-naïve patients consist of two nucleoside/nucleotide analogue reverse transcriptase inhibitors (NRTI) plus either efavirenz (EFV) (a nonnucleoside reverse transcriptase inhibitor (NNRTI)), or one of several boosted protease inhibitors (PIs), or an integrase strand transfer inhibitor.[5] If a preferred regimen fails, there are numerous other drugs that can be used in a variety of possible combinations. Continued suppression of HIV-RNA can be maintained indefinitely in the majority of individuals who adhere to appropriate HAART regimens.

Regulatory History of Antiretroviral Drug Development and Approval

Most antiretroviral drugs initially entered the market via accelerated approval based on changes in surrogate endpoints, primarily plasma HIV-RNA levels but also $CD4^+$ cell counts, before routine monitoring with HIV-RNA. Before 1997, traditional approvals were based on clinical endpoint trials assessing the effects of a drug on mortality and/or HIV disease. With the success of combination therapy, subsequent decline of HIV-related illnesses (Palella et al. 1998; Hogg et al. 1999), and the routine use of HIV-RNA monitoring to assess response to treatment, it became clear that a requirement for clinical endpoint trials for every traditional approval was no longer feasible. In July 1997, we 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.[6]

[5] Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents, Department of Health and Human Services Panel on Antiretroviral Guidelines for Adults and Adolescents — A Working Group of the Office of AIDS Research Advisory Council (http://aidsinfo.nih.gov/contentfiles/AdultandAdolescentGL.pdf).

[6] See http://www.fda.gov/forconsumers/byaudience/forpatientadvocates/hivandaidsactivities/ucm117940.htm#endpoints.

In 1996 and 1997, a collaborative group of pharmaceutical, academic, and government scientists investigated relationships between treatment-induced changes in HIV-RNA levels and clinical endpoints collected from ongoing and completed antiretroviral trials (Murray et al. 1999; Hill et al. 1999). Several analyses of more than 5,000 patients in multiple trials identified a relationship between initial decreases in plasma 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 $CD4^+$ cell counts and HIV-RNA levels, prior drug experience, and treatment regimen (Marschner et al. 1998).

Based on these data, the Antiviral Drug Advisory Committee concluded 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. Specifically, the committee stated 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), a surrogate endpoint "reasonably likely to produce long-term benefits," 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), a surrogate endpoint more convincingly related to long-term benefit in the setting of life long therapy. The committee also recommended that changes in $CD4^+$ cell counts be consistent with observed HIV-RNA changes when considering approval of an antiretroviral drug.

Subsequently, additional data further supported the utility of an endpoint of viral load suppression for predicting a clinical benefit in HIV progression. Such data include:

- Analysis of 12 clinical endpoint trials (originally submitted to the FDA in support of approval) that showed that a 0.5 log reduction in HIV-RNA between treatment arms was also associated with a reduction in clinical disease progression

- Results from the Strategies for Management of Anti-Retroviral Therapy (SMART) trial that showed that a strategy of continuous viral suppression provided a lower risk of disease progression than a strategy of drug conservation that allowed for treatment holidays until $CD4^+$ cell counts declined to a specified amount (SMART Study Group 2006)

- Epidemiologic reports (Hogg et al. 1999) that showed that the current treatment strategy of maximal viral suppression with HAART has dramatically reduced AIDS morbidity and mortality

- Data from numerous trials that showed incomplete viral suppression results in emergence of viral resistance, viral rebound, and loss of efficacy of individual drugs and sometimes entire drug classes

All drugs that received accelerated approval, either before 1997 or since that time, subsequently received traditional approval. Since 1997, 13 antiretroviral drugs entered the market via an accelerated approval based on 24-week changes in viral load. All of these drugs were confirmed to have durable virologic suppression at 48 weeks and beyond. Although a percentage of people on HAART develop virologic failure over time, in no case did longer term data reveal that a drug lost the substantial efficacy initially seen at time of accelerated approval. However, longer term data have shown more subtle differences between treatment arms comparing different drugs or dosing regimens and have been useful for choosing optimal doses or preferred regimens in treatment guidelines.

Given that HIV-RNA is a validated surrogate for predicting efficacy of antiretrovirals, a paradigm of accelerated approval (based on viral load changes at 24 weeks) followed by traditional approval (based on viral load changes at 48 weeks) is no longer needed for the development of antiretrovirals. Instead traditional approval can be the initial approval for all antiretroviral drugs, with the duration of viral load assessments dependent on the population studied, as will be described in this guidance. Table 1 summarizes recommended treatment durations to support approvals of indications for the listed subgroups.

Table 1: Recommendations for Efficacy and Safety Determination Time Points According to HIV Patient Population

Patient Population	**Efficacy Determination Time Point**	**Safety Determination Time Point**
Treatment-naïve or limited[a] previous treatment	Virologic response at 48 weeks	Safety outcomes through 48 weeks
Treatment-experienced with remaining options	Virologic response at 24-48 weeks[b]	Safety outcomes through 24-48 weeks
Treatment-experienced with no or few remaining options	Virologic response at 2 weeks plus virologic follow-up at 24 weeks	Safety outcomes through 24 weeks

[a] Previous treatment with first regimen with no documented virologic failure.

[b] Twenty-four weeks of data is appropriate for drugs that have some benefit over existing options (e.g., better efficacy, tolerability, ease of administration), while 48 weeks is recommended for drugs with comparable characteristics to existing options.

III. DEVELOPMENT PROGRAM

A. General Considerations

1. Pharmacology/Toxicology Development Considerations

Pharmacology/toxicology development for HIV-1 antivirals should follow existing guidances for drug development.[7]

The above-referenced guidances suggest that nonclinical combination studies generally should be conducted to support clinical trials for combination drugs involving two entities in early stages of development. In the ICH guidance for industry *M3(R2) Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals*, section I.C., Scope of the Guidance, states, "Pharmaceuticals under development for indications in life-threatening or serious diseases (e.g., advanced cancer, resistant HIV infection, and congenital enzyme deficiency diseases) without current effective therapy also warrant a case-by-case approach to both the toxicological evaluation and clinical development in order to optimize and expedite drug development."

For new HIV drug combinations of early stage entities that are not expected to offer benefits over currently effective therapy, combination toxicology studies usually should precede combination clinical trials. However, usually no more than two drugs should be tested simultaneously in a particular arm of a toxicology study. The design of such studies should be discussed with the DAVP. For combinations that are expected to offer benefits over currently effective therapy such as treating drug-resistant HIV in patients with few remaining options, combination toxicology studies may not be warranted when all of the following apply:

- Mechanisms of action or in vitro data of potential off-target effects of the individual drugs do not suggest a potential for additive or synergistic toxicity.

- Studies in animals or humans of absorption, distribution, metabolism, and excretion of the individual drugs do not suggest potential for an unmanageable interaction (one that cannot be addressed with dose adjustments) or serious toxicity for the combination.

- Toxicology studies (of at least 3 months duration) of the individual drugs show a substantial safety margin for the intended clinical dose(s) or exposures.

- Phase 1 clinical data in healthy volunteers or HIV-infected patients receiving the individuals drugs show no substantial or unmanageable safety concerns. Phase 1 data should include single- and multiple-dose PK and safety trials, at a minimum.

[7] See the ICH guidances for industry *M3(R2) Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals* and *S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals*.

Additional safety data from phase 1 and phase 2 trials are encouraged and may be warranted if one or more of the drugs demonstrate a potential serous safety risk.

- There are no concerning overlapping toxicities for the individual drugs based on animal toxicology studies and phase 1 or phase 2 clinical data.

- Clinically significant PK drug-drug interactions are considered unlikely or can be reliably managed with dose adjustments such that safety margins based on individual drug exposures are not exceeded.

After considering the previous points, sponsors can first evaluate (in phase 1 and phase 2 trials) in HIV-infected patients who are treatment-naïve or have remaining treatment options, drug combinations intended to treat drug-resistant HIV. After initial trials in treatment-naïve patients or patients with several available treatment regimens have helped to define the most active doses, patients with few or no remaining treatment options can be studied. This approach helps to ensure that patients with no remaining treatment options are not exposed to suboptimal doses or combinations that could severely jeopardize their chance (perhaps only chance) for achieving durable virologic suppression. However, combination trials in healthy volunteers or healthy HIV-infected patients should not be the first-in-human trials unless the drugs cannot be administered separately and unless combination toxicology studies have been completed according to ICH guidance.

Nonclinical combination studies of an investigational antiretroviral plus an approved antiretroviral generally are not warranted and are not feasible because individual antiretrovirals are often combined with multiple other antiretrovirals in multiple different regimens over a lifetime of treatment. Therefore, unless data from nonclinical studies of an investigational antiretroviral suggest a potential for serious synergistic toxicity with an approved therapeutic drug combination, toxicology studies are not expected.

Applicants can choose to submit carcinogenicity studies with an initial new drug application (NDA) or as required postmarketing studies.

2. Nonclinical Virology Development Considerations

Antiretrovirals for the treatment of HIV-1 should be tested in cell culture for antiviral activity before submission of an initial investigational new drug application (IND). Information about pre-investigational new drug applications and information regarding appropriate nonclinical assays is available from the FDA.[8] Additional recommendations for general antiviral drug development can be found in the guidance for industry *Antiviral Product Development — Conducting and Submitting Virology Studies to the Agency*.

[8] See the FDA Web site http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/InvestigationalNewDrugINDApplication/Overview/ucm077546.htm.

a. Mechanism of action

The mechanism by which an antiretroviral drug specifically inhibits HIV replication or a virus-specific function should be investigated in studies that include evaluation of the effect of the drug on relevant stages of the virus life cycle. Mechanism of action investigations should include appropriate controls for assessing the specificity of anti-HIV activity, which may include assessments of activity against HIV proteins that are not targeted by the candidate drug, relevant host proteins, and other viruses.

b. Antiviral activity in cell culture

The antiviral activity of a new drug should be characterized in cell culture to demonstrate anti-HIV activity and identify a target plasma concentration for evaluation in HIV-infected patients. Anti-HIV activity studies should include assessments against a broad range of clinical and laboratory viral isolates including different groups and subtypes (or clades). The effective concentration at which virus replication is inhibited by 50 and 90 percent (e.g., EC_{50} and EC_{90} for cell-based assays; IC_{50} and IC_{90} for biochemical or subcellular assays) should be determined using a quantitative assay.

Sequestration of the drug by serum proteins also should be assessed and a serum-adjusted EC_{50} value determined. We recommend evaluation of the drug's antiviral activity at different concentrations of human serum and extrapolation to a 100 percent human serum EC_{50} value.

c. Cytotoxicity

The cytotoxic effects of the drug should be quantified directly in the cells used for assessing anti-HIV activity, and a 50 percent cytotoxic concentration (CC_{50}) and a therapeutic index should be calculated. Cytotoxicity also should be assessed using various cell lines and primary cells cultured under proliferating and nonproliferating conditions. Cytotoxicity and mitochondrial toxicity assessments under proliferating conditions should be evaluated with drug exposures for several divisions.

d. Combination antiviral activity

We anticipate that most, if not all, antiretrovirals will be used to treat HIV-1 in combination with other approved drugs. Early in development, cell culture combination antiviral activity relationships of the new drug with two representatives of each antiretroviral drug class should be evaluated to determine whether the combination antiviral activity is antagonistic. If antagonism is seen with either member of a class, all members of the class should be evaluated. Additional combination antiviral activity studies with other candidate antiretroviral drugs should be conducted if future combination therapy with other drugs is anticipated. For all combination antiviral activity assessments, sponsors should provide combination index values when the two drugs are combined at or near their individual EC_{50} values, and studies should include controls for cytotoxicity. Combination antiviral activity relationships for HIV and

hepatitis C virus (HCV) or hepatitis B virus (HBV) drugs with similar mechanisms of action (e.g., nucleo(t)side analogue polymerase/reverse transcriptase inhibitors, PIs) also should be assessed before testing combinations of the drugs in HIV/HCV or HIV/HBV co-infected patients.

e. Activity in animal models

Demonstration of anti-HIV activity in an animal model is not needed.

f. Resistance and cross-resistance

The ability of HIV to develop resistance to an antiretroviral when subjected to drug pressure should be examined in appropriate cell culture models. Amino acid substitutions associated with the development of resistance to the candidate drug should be determined and validated by introducing the mutations into the HIV genome, and determining the conferred fold-shift in susceptibility using appropriate cell culture and/or biochemical assays. Results from these studies should be used to: (1) identify resistance pathways; (2) determine whether the genetic barrier for resistance development is high or low; (3) predict whether the genetic barrier for resistance may vary as a function of concentration of the new drug; (4) assess the potential for cross-resistance with other anti-HIV drugs; and (5) support the drug's hypothesized mechanism of action.

Resistance studies should include evaluation of the potential for cross-resistance, both to approved drugs and also to drugs in development when possible, particularly focusing on those in the same drug class and other classes targeting the same protein or protein complex. The antiviral activity of the investigational drug should be assessed against mutant viruses that are resistant to drugs within the same drug class as the investigational drug as well as a representative sample of viruses resistant to other approved antiretroviral drugs.

3. Drug Development Population

We encourage the evaluation of antiretroviral drugs in a wide range of patients including treatment-naïve and treatment-experienced patients, as appropriate. However, the drug development population depends to a large extent on specific characteristics of the drug such as resistance profile, tolerability, pharmacologic profile, and route of administration. A drug with a daily subcutaneous or intravenous route of administration may be acceptable for a highly treatment-experienced patient with few remaining options, but generally would not be considered appropriate for a treatment-naïve individual. A drug with a favorable resistance profile that retains activity to viral strains resistant to approved drugs is likely to fill an unmet medical need in treatment-experienced patients. However, such a drug need not be restricted to treatment-experienced patients if it is well tolerated and favorable in other aspects (e.g., convenient dosing schedule). Investigational drugs intended for treatment-naïve patients should be at least as efficacious, well tolerated, and convenient to administer as approved drugs for use in

treatment-naïve patients and ideally should have some favorable characteristic for at least a subgroup of naïve patients if deficient in another aspect.

We encourage the study of antiretrovirals in patients having the greatest need for new drugs, such as patients who cannot tolerate other antiretrovirals or have developed resistance to multiple antiretrovirals. We realize that trials in heavily treatment-experienced patients may need to be supported by preliminary data from trials in healthy volunteers and in HIV-infected populations with less or no prior antiretroviral therapy to define preliminary activity, safety, and pharmacokinetics (e.g., drug-drug interaction trials).

HIV is a disease that is present worldwide and clinical trials typically are conducted internationally. However, trials should include adequate U.S. patient representation and patients infected with Clade B virus to ensure applicability of trial results to the U.S. population. An adequate representation of males and females, races, ages, and weights are recommended during all stages of drug development, especially in phase 3 trials. Inclusion of a diverse patient population early in drug development may help to identify potential efficacy or safety issues and can help to inform the design of phase 3 trials. Sponsors should share with the FDA their pretrial initiation work to ensure the sites selected have sufficient numbers of women and racial representation to enroll in phase 2 and 3 clinical trials.

4. Early Phase Clinical Development Considerations

a. First-in-human trials

For first-in-human trials, we recommend single- and multiple-ascending-dose trials in healthy adult subjects to assess safety and pharmacokinetics and to avoid development of resistance that could occur from subtherapeutic exposure in HIV-infected individuals.

b. Phase 1b (proof-of-concept) trials

The first proof-of-concept trial in HIV-infected patients should be a multiple-dose study that allows for short-term (e.g., several days to 2 weeks depending on the drug class and resistance profile in cell culture) evaluation of a drug's effect on reducing HIV-RNA levels from baseline and also provides for evaluation of safety for a short duration. Duration of monotherapy should be minimized to reduce the risk of resistance while still being able to assess activity. Mean changes in HIV-RNA from baseline should be the primary endpoint. Examples of proof-of-concept studies include:

- A randomized placebo-controlled trial comparing the new investigational drug, at several dose levels, to placebo in HIV-infected patients who are treatment-naïve or who are not currently receiving therapy but who had limited exposure to therapy in the past. The trial duration depends on the anticipated resistance barrier of the drug based on cell culture studies. Some drugs with an anticipated low genetic barrier to resistance would not be appropriate candidates for study in

a monotherapy trial of any duration. Drugs with a higher barrier to resistance emergence can be studied for up to 2 weeks.

- A randomized placebo-controlled trial comparing the new investigational drug, at several dose levels, to placebo in HIV-infected patients who are currently receiving HIV treatment with approved drugs but have not achieved or maintained viral suppression on their current regimen. Adding one new drug to a regimen not producing complete viral suppression is sometimes referred to as *functional monotherapy*. Functional monotherapy is not recommended for long durations. The primary assessment of activity should occur at 2 weeks (or perhaps sooner for some drugs). After the initial placebo-controlled comparison of efficacy, patients can be followed on open treatment for longer periods for safety, durability of response, and emergence of resistance. However, we recommend that trials contain provisions for changing the background regimen after 2 weeks in an attempt to maximize the likelihood of a fully suppressive regimen. Also, patients randomized to placebo can be allowed to receive the new investigational drug after 2 weeks in addition to an optimized background regimen, provided that there are supporting pharmacology/toxicology data for longer term administration.

c. Phase 2 trials and dose finding

The goal of early phase 2 trials is to characterize an active, tolerable, and safe dose(s) of an antiretroviral drug as part of a combination regimen for further study in phase 3 trials. Sponsors should conduct mechanistic modeling of the concentration-viral kinetics and the concentration-safety profile from short-term monotherapy trials to choose doses for early phase 2 trials. As a general rule, doses selected for phase 2 should provide exposures expected to exceed, by several-fold, the protein binding-adjusted, cell culture EC_{50} value of the drug for the relevant HIV genotype/subtype. However, for some drug classes, specifically NRTIs, intracellular triphosphate concentrations are more related to pharmacodynamic effect than plasma concentrations. Sponsors should avoid selecting doses that provide exposures that are expected to be largely subtherapeutic to reduce the risk of selecting for resistant virus.

Phase 2 dose-ranging studies that have demonstrated a significant dose response can provide supportive data for an approval of an antiretroviral drug. Generally, dose-comparison studies should include a large enough range of doses to demonstrate a dose- or exposure-response relationship.

5. Efficacy Considerations

In general, NDAs should include at least two adequate and well-controlled trials conducted in the proposed population(s) intended for labeling. Applicants can submit an NDA in a single population, either treatment-naïve or treatment-experienced patients. Alternatively, applicants can choose to pursue an indication for both treatment-naïve and -experienced patients. In this circumstance, the NDA should contain at least one

adequate and well-controlled phase 3 trial in each patient population, with adequate supporting data from phase 2 trials. Sponsors should consult existing guidance regarding circumstances in which one phase 3 clinical trial may be supportive of approval.[9]

Applicants should consult 21 CFR 300.50 for specific regulatory considerations regarding fixed-dose combinations. In brief, two or more drugs may be combined in a single dosage form when each component makes a contribution to the claimed effects of the drug, and the dosage of each component is such that the combination is safe and effective for a significant patient population requiring such concurrent therapy as defined in the labeling for the drug.

HIV treatment development plans may be eligible for consideration under 21 CFR part 312, subpart E, Drugs Intended to Treat Life-Threatening and Severely-Debilitating Illnesses, fast track, breakthrough therapy designation,[10,11] or priority review if the specifics of the development plan justify such an approach.

6. Safety Considerations

The majority of antiretroviral approvals were based on databases including approximately 500 patients receiving the approved dose for at least 24 to 48 weeks depending on the population. For indications in treatment-naïve patients or patients with limited prior treatment experience, applications should include at least 500 individuals receiving the intended dose for 48 weeks duration. For heavily treatment-experienced patients, safety data on 300 to 500 patients receiving the intended dose for 24 weeks should be sufficient. For indications in patients with intermediate levels of treatment experience, 500 patients for 24 to 48 weeks may be appropriate, depending on the particular drug's efficacy or advantages over other available treatment options. Applicants are encouraged to discuss their proposed safety database with the DAVP before submitting an NDA. On occasion, specific findings in nonclinical or phase 1 and phase 2 development may indicate the need for a database that is larger or longer in duration to adequately evaluate potential drug toxicity.

Applicants should provide controlled and comparative safety data. Safety data from uncontrolled protocols or treatment 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 to assess in uncontrolled safety data. Trials assessing dose response are often particularly useful for evaluating drug-related adverse reactions.

---

[9] See the guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products.*

[10] See section 506 of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 356) as amended by section 902 of the Food and Drug Administration Safety and Innovation Act of 2012.

[11] See the FDA Web site Fact Sheet: Breakthrough Therapies at http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/SignificantAmendmentstotheFDCAct/FDASIA/ucm329491.htm.

B. Specific Efficacy Trial Design Considerations

1. Trial Design and Trial Population

The appropriate trial design depends on the population being studied. HIV-infected populations typically studied include:

- Treatment-naïve
- Treatment-experienced with available approved treatment options
- Treatment-experienced with few or no available approved options

It is important to emphasize that treatment-naïve patients have several approved treatment options that are highly effective, tolerated, and convenient to use (e.g., 1 tablet or capsule once daily for an entire regimen). Although an active and tolerable antiretroviral regimen can be identified in 24 weeks or less, modest differences in virologic efficacy, emergence of resistance, and tolerability are sometimes detected when treatment-naïve patients are followed through 48 weeks and beyond. Given that the initial regimen usually is the best and preferred regimen and that loss of response to an initial regimen can often affect the choice of subsequent drugs because of resistance, regimens for treatment-naïve patients are evaluated stringently and are compared to known, high-performing, control regimens.

Lower efficacy or tolerability of a new drug/regimen compared to known controls in treatment-naïve patients is an important issue that can affect approval for this use or lead to precautionary language in labeling. Standard regimens for treatment-experienced patients are less well defined than for treatment-naïve patients; it is sometimes appropriate to evaluate the effectiveness of potentially promising drugs in combination with individualized background drugs for treatment-experienced patients at time points earlier than 48 weeks.

Treatment-Naïve Patients

In treatment-naïve patients, who cannot be denied active treatment, the most feasible trial design is a randomized active-controlled noninferiority trial (see Appendix B for a discussion of noninferiority margins). In this design, patients will be randomized to a standard three-drug regimen or the same standard regimen with the investigational drug substituting one of the components of the regimen and followed for at least 48 weeks.

Multiple doses of the investigational drug can be studied in active-controlled noninferiority studies to better define an optimal dose (but a dose known to be less effective could not be ethically chosen). An observed dose response would strongly support efficacy.

Add-on superiority trials (e.g., three approved drugs plus the investigational drug compared to three approved drugs) are considered less feasible because the response rate

in treatment-naïve patients is high (greater than 80 percent); lack of response often occurs for reasons other than virologic failure, such as poor adherence or early drop out for adverse events. Consequently, four active drugs often have not shown improved efficacy over three drugs in this population. In addition, showing superiority to current commonly used control regimens in an active-controlled substitution trial is difficult for the same reasons as described above.

Treatment-Experienced Patients With Available Treatment Options

An active-controlled noninferiority comparison (as described above) with or without comparisons of multiple doses of the investigational drug is an acceptable trial design. For this population, patients should be followed for at least 24 to 48 weeks. NDA submissions can be made after an analysis at 24 weeks, if the drug demonstrates superiority over approved drugs. Choice of the active control and control arm regimen is less straightforward than treatment-naïve trials because second-line regimens are not well defined in treatment guidelines and generally are left up to clinical judgment depending on the situation. However, we recommend using controls and control arm regimens that were previously studied in large randomized trials to justify the choice of a noninferiority margin (see Appendix B).

Add-on superiority trials where patients are randomized to a new regimen consisting of approved drugs versus a new regimen of approved drugs plus the investigational drug is another possible trial design. The approved drugs in the regimen usually are selected after taking into account patient history and resistance testing. It is desirable for patients on both arms to have a sufficient number of drugs to construct a fully suppressive regimen. However, if the enrolled patient population has too many remaining treatment options, particularly drugs with a high level of potency, it is likely that adding another drug to the regimen would not demonstrate superiority.

If two new investigational drugs are available for study at the same time, a randomized controlled superiority trial with a factorial-type design can be used. This design may be useful when studying patients who are unable to construct a viable antiretroviral regimen from approved drugs. In this type of trial design, where both A and B are investigational drugs, patients could be randomized to one of the following trial arms:

- Arm 1: Approved drugs + A+ B
- Arm 2: Approved drugs + A
- Arm 3: Approved drugs + B

A fourth arm, of only approved drugs, could be considered if patients have enough remaining approved drugs to construct a regimen, but if this were the case, showing superiority of adding drugs to the regimen may be difficult. To demonstrate efficacy for drug A, arm 1 would need to be superior to arm 3, and to demonstrate efficacy of drug B, arm 1 would need to be superior to arm 2.

Treatment-Experienced Patients With Few or No Available Approved Treatment Options

This population is also referred to as heavily treatment-experienced. Noninferiority studies generally are not feasible in this population because there usually is no appropriate active control with a sufficiently well-described effect that can be used to define a noninferiority margin.

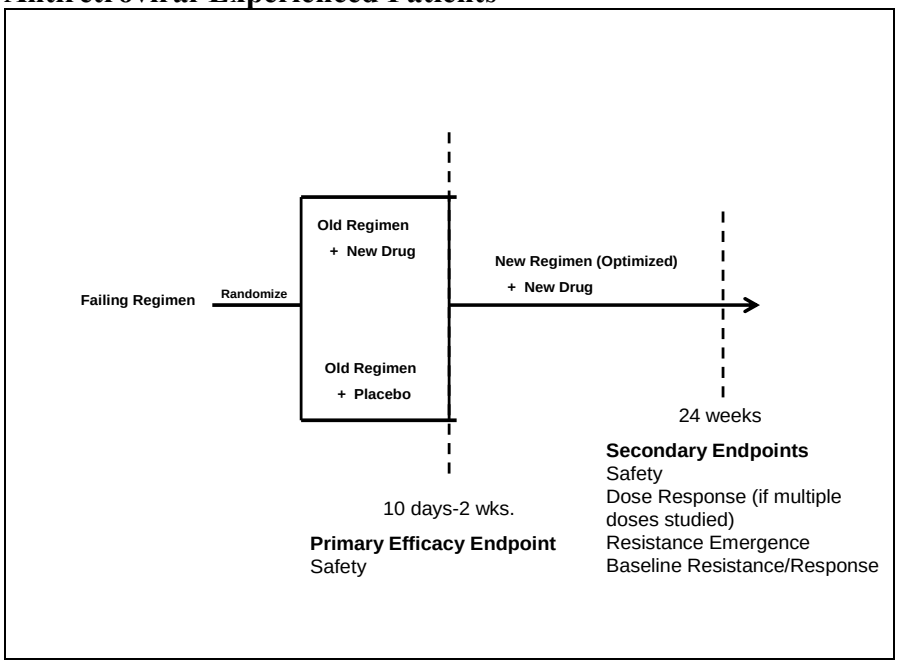
If two investigational drugs with activity against multidrug resistant virus are available for study simultaneously, the factorial design as described above is a reasonable option.

When only one new drug is available for study in a clinical trial, a randomized placebo-controlled superiority trial should be conducted where the primary endpoint is assessed at an early time point (see Figure 1). Longer term placebo-controlled comparisons have fallen out of favor because they run the risk of emergence of resistance to the investigational drug or the background drugs. In our recommended design, patients experiencing ongoing viral replication on their current regimen and who need a new drug to construct a new viable regimen are continued on their current regimen, and randomized to add either placebo or the new investigational drug (randomization to the investigational drug could be for one or more dose levels). The primary efficacy evaluation of investigational drug versus placebo occurs over a short duration (7 days to 2 weeks), before development of a significant risk for resistance to the new drug or additional resistance to the background drugs. After the placebo comparison, all patients can receive the investigational new drug (at one or various dose levels) added to a new background of approved drugs that are optimized by resistance testing. In this proposal, a second assessment occurs at 24 weeks to assess for:

- A dose response (if multiple doses are included)
- Response by baseline susceptibility or resistance profile
- Safety
- Durability of initial response
- Emergence of resistance to the investigational drug and other drugs in the regimen

The primary efficacy analysis is the short duration (e.g., 2 weeks) comparison to placebo. At 24 weeks, the comparison is no longer controlled unless a dose response is being evaluated. Given that doses chosen for study in HIV trials usually are on the plateau portion of a dose-response curve, demonstration of a dose response is considered unlikely. This design is similar to one of the recommended phase 1b trial designs discussed above, except that this phase 3 trial is larger and allows for a more thorough evaluation of baseline characteristics and response at 24 weeks. In addition this trial should be conducted after smaller initial proof-of-concept trials identify reasonably active doses to reduce the likelihood of administering suboptimal doses to this vulnerable population. Evaluation for both safety and efficacy beyond 24 weeks is recommended and could be accomplished during the postmarketing period.

Figure 1: Schematic of Possible Trial Design in Heavily Antiretroviral-Experienced Patients



This type of study design, which includes a primary efficacy analysis at 2 weeks (or less) and a safety analyses at 24 weeks, may be appropriate for a population of heavily treatment-experienced patients when the investigational drug is expected to offer antiviral activity in the setting of multiple-drug resistance. First drugs of a new class or *second generation* drugs of an existing class that can treat drug-resistant strains are candidates for this type of study design. Trials conducted in this population would support only a limited treatment indication for use in patients who cannot construct a viable regimen without a new antiretroviral drug.

Criticisms of this approach primarily relate to the uncontrolled design of the study beyond the primary 2-week comparison and the concern that it doesn't allow for an adequate assessment of virologic durability or safety. However, the unmet medical need in this population and the potential to decrease further development of resistance in the background regimen of trial patients outweigh any modest loss of certainty in the interpretation of results from this type of trial design.

After decades of antiretroviral drug development, many experts agree that active antiretroviral drugs can be identified within days to weeks of antiviral load monitoring based on early viral load kinetics. Durability of response is related to the ability to use a drug with an active supportive regimen. In fact, even drugs with low barriers of resistance have become *preferred* when combined with other active drugs in treatment-naïve patients. In a heavily treatment-experienced population, multiple types of regimens likely will be used with a new drug, so there is no well-defined benchmark to compare noninferiority. The assessments that the above trial design provides — with respect to comparative short-term activity, longer term observations for virologic rebound or virologic durability, and safety and potential dose-response — are adequate to support

approval of a limited indication for a population at high risk of suffering substantial HIV-related complications.

2. Randomization, Stratification, and Blinding

We encourage sponsors to conduct double-blind trials whenever feasible. For add-on superiority trials of a new antiretroviral plus background therapy compared to background therapy alone, patients randomized to the latter should receive a matching placebo. In open-label protocols, patients may be more likely to drop out of the trial if they know they are not receiving the new treatment.

There are situations in which blinding drugs or regimens may not be feasible, but 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 drug interactions with other drugs in the regimen; however, this could be accomplished by similarly dose adjusting the placebo. In studies adding test drugs to a common background in most cases blinding only one component of a regimen is needed. Background therapy does not need to be blinded.

Sponsors designing studies in which blinding may be difficult or infeasible should discuss the proposal with the DAVP in advance to review potential modifications that might facilitate blinding and to discuss the potential effect 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 because 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 we anticipate additional sensitivity analyses using different methods of handling treatment discontinuations or missing data.

Sponsors should consider stratification of patients by important baseline factors such as viral load (less than 100,000 copies/mL versus greater than or equal to 100,000 copies/mL), CD4 cell count (less than 200 versus greater than or equal to 200), and geographic area. Baseline resistance scores (phenotypic, genotypic, or overall susceptibility) can be used as a stratification factor in treatment-experienced trials.

3. Choice of Controls

Sponsors should include treatment regimens consistent with standards of clinical practice while the trial is being conducted. Because of the evolving nature of accepted standards of HIV treatment, appropriate comparison regimens can be expected to change over time. In general, current HIV treatment guidelines emphasize the importance of using at least three potentially active drugs (if possible) when constructing a regimen. However, some of the newer approved drugs have potency that could possibly support study of two-drug combinations. From a patient management perspective, use of control regimens that have

been determined to be suboptimal, as based on clinical studies or consensus of expert panels reviewing pertinent data, would jeopardize the viability of a trial and possibly future treatment options for patients, and therefore should not be used. Protocol proposals with control arms that deviate from current standards of care should be discussed with the DAVP before implementation and may require ethics consultation.

Cross-class comparisons may be appropriate for treatment-naïve trials. An investigational drug with the potency of an NNRTI, integrase inhibitor, or boosted PI can be compared to EFV, an integrase inhibitor, or one of the preferred boosted PIs. If two naïve studies are being conducted, an in-class comparison and a cross-class comparison trial can provide useful comparative information for a prescriber. In particular, the value of EFV as a comparator in active-controlled trials in treatment-naïve patients is: (1) it has been used in many trials as a control arm for historical reference; (2) its efficacy has not been substantially exceeded by other newer drugs; (3) the choice of noninferiority margin is clear (see Appendix B); and (4) it has wide acceptance among clinicians.

For treatment-naïve trials, a drug with the potency of a nucleo(t)side reverse transcriptase inhibitor can be compared to one of the other two NRTIs in the regimen. In current preferred regimens the active comparator can be tenofovir, lamivudine, or emtricitabine. The value of using one of these drugs as comparators is: (1) they have been used in many trials as controls so they provide historical reference; and (2) they have wide acceptance among clinicians. When studying an NRTI in a noninferiority study, the third drug should be EFV or another similar NNRTI and not a boosted PI. The relative contributions of NRTIs to an EFV-based regimen can be reasonably inferred from previous data. This is not the case for regimens that include boosted PIs. See Appendix B for the recommended noninferiority margin for a noninferiority trial that uses EFV as the active control.

For treatment-experienced patients, there are no clear standard regimens. Active controls depend on the exact patient population studied with respect to baseline resistance and also depend on a sufficiently robust demonstration of efficacy of active controls in previously conducted trials. Noninferiority margins can be based on a rationale similar to that described in Appendix B. Noninferiority trial proposals should be discussed with the DAVP in advance.

4. Efficacy Endpoints

We recommend the following primary efficacy endpoints for phase 2 and 3 studies:

- **For treatment-naïve trials:** the proportion of patients with HIV-RNA levels below the limit of assay detection at 48 weeks using a sensitive, FDA-licensed test. The method for calculating these proportions is described in Appendix A.

- **For trials in treatment-experienced patients with multiple remaining approved drug options:** the proportion of patients with HIV-RNA levels below the limit of assay detection at 48 weeks using a sensitive, FDA-licensed test. A

24-week time point can be used for superiority comparisons when a drug is expected to offer an advantage over currently available options.

- **For trials in treatment-experienced patients with few remaining approved options:** the proportion of patients with HIV-RNA decreases from baseline exceeding 0.5 log at an early time point (approximately 2 weeks).

Secondary endpoints should include:

- Mean changes in viral load from baseline for treatment-experienced patients
- Changes in CD4 cell counts from baseline

5. Trial Procedures and Timing of Assessments

Recommended critical time points for measuring viral RNA depend on the patient population studied. Early time points (1 to 4 weeks) are critical assessments for heavily treatment-experienced patients. Beyond the first month, HIV-RNA, $CD4^+$ cell counts, and safety assessments are typically collected at weeks 8, 12, 16, 24, 36, and 48 and every 3 to 6 months beyond 48 weeks. Longer term follow-up out to 96 weeks and beyond is recommended particularly for treatment-naïve patients. Longer term follow-up can be completed as a postmarketing commitment or a postmarketing requirement if there is a safety concern identified in the 48-week dataset that needs further evaluation.

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. Statistical Considerations

Sponsors should designate the hypotheses to be tested before trial initiation. These hypotheses should be stated in the protocol or the statistical analysis plan (SAP). If sponsors choose to test multiple hypotheses, they should address issues related to the potential inflation of false positive results (overall type I error rate) caused by multiple comparisons. These issues should be discussed with the DAVP in advance of trial enrollment, and should be incorporated into SAPs as appropriate.

a. Analysis populations

The following definitions apply to various populations for analyses in HIV clinical trials:

- **All randomized (AR) population** — All patients who are randomized. This population is sometimes referred to as the intent-to-treat population.

- **All treated population** — All patients who are randomized and receive at least one dose of assigned therapy during the trial. This population is sometimes referred to as the safety population or the modified intent-to-treat population.

b. Efficacy analyses

In treatment-naïve trials and trials in treatment-experienced patients with multiple remaining approved drug options, the primary efficacy endpoint should be the proportion of patients with HIV-RNA below the limit of assay detection at 48 weeks (or 24 weeks for drugs with a likely treatment advantage over available options for treatment-experienced patients) using a sensitive, FDA-approved viral load assay. The method for calculating the proportion is described in Appendix A.

The primary efficacy analysis should be adjusted for at least one or two of the most important covariates (e.g., baseline HIV-RNA). The covariates that will be included in the primary analysis should be prespecified in the protocol. Cochran-Mantel-Haenszel analyses and Breslow-Day statistics can be used to examine the homogeneity of treatment effects. The calculation of the difference between two proportions and its confidence interval can be based on stratum-adjusted Mantel-Haenszel proportions.

For subgroup analyses, the analysis of the primary efficacy endpoint should be performed within important demographic and baseline characteristics such as sex, race, age group, region, baseline HIV-RNA viral load, baseline $CD4^+$ cell count, clade, and baseline resistance score. The purpose of the subgroup analyses is to evaluate the consistency of the primary efficacy endpoint result across these subgroups. It is important to recognize, however, that simply by chance a drug that has a homogeneous overall effect in a trial population will often show different effects in some subgroups, sometimes even showing significant heterogeneity, in any given trial. Therefore, such subgroup results should be interpreted with caution.

We encourage sponsors to collect the data regarding drug-adherence and change of treatment including switching treatment and adding the additional therapy. These data are particularly important to confirm and determine the reasons for discontinuation among the patients who discontinue the assigned therapy early so that these patients can be appropriately classified in the analysis.

c. Noninferiority margin

In noninferiority trials, the choice of noninferiority margin for statistical hypotheses should be discussed with the DAVP before study initiation because one margin is not appropriate for all study designs. The sponsor should attempt to define a margin (M_1) 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 (see Appendix B).

In addition, the noninferiority margin (M_2) should be smaller than M_1 to preserve a clinically important effect compared to an active control. For noninferiority testing, sponsors should employ two-sided 95 percent confidence intervals adjusted for multiple comparisons or other appropriate testing procedures. Both noninferiority and superiority can be assessed in a noninferiority study provided that the noninferiority comparison is conducted first and superiority is conducted only after noninferiority is met, and choice of delta has been specified before study initiation and/or provided so that the choice of delta can be justified based on previous clinical data. For additional information regarding noninferiority studies in general, see Appendix B, ICH E10, and the draft guidance for industry *Non-Inferiority Clinical Trials*.[12]

d. Missing data

There is no single optimal way to deal with missing data from clinical trials. Sponsors should make every attempt to limit loss of patients from the trial, and when the loss is unavoidable, collect information that can help explain the cause of the loss and the final status of the patient. Analyses excluding patients with missing data or other post-treatment outcomes are potentially biased because patients who do not complete the trial may differ substantially in both measured and unmeasured ways from patients who remain in the trial. The method of how missing data will be handled should be specified in the protocol or the SAP. A patient retention and follow-up plan should be included in the protocol providing details on how to minimize missing data and collect follow-up information.

e. Interim analyses and data monitoring committees

If interim (or futility) analyses are performed, these analyses should be prespecified in the protocol and the SAP. The purpose of the interim analysis should be stated in the analysis. If an adaptive design such as withdrawal of a treatment arm or sample size re-estimation based on an interim analysis is applied, then the adaptive design procedures should be prospectively prespecified.[13] It is important that the interim analysis does not affect study conduct and thereby compromise trial results.

Use of a data monitoring committee (DMC) may be appropriate depending on the design of the proposed phase 3 trial. If a DMC is used, a detailed charter with the composition

[12] When final, this guidance will represent the FDA's current thinking on this topic. For the most recent version of a guidance, check the FDA Drugs guidance Web page at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

[13] See ICH E9 and the draft guidance for industry *Adaptive Design Clinical Trials for Drugs and Biologics* (when final, this guidance will represent the FDA's current thinking on this topic).

of the committee members and the operational procedures should be provided for review.[14]

f. Other analyses of interest and secondary endpoints

Sponsors can present secondary analyses on other endpoints of interest. An analysis of change in CD4 cell count from baseline at Week 24 or 48 between the treatment groups is a recommended secondary endpoint. In the event that a CD4 cell count at Week 48 time window is missing, we suggest that there be a planned analytic approach to impute missing data. Examples include, but are not limited to, last observation carried forward, baseline observation carried forward, and mixed-effect models. It may be useful to compare results with other approaches to examine sensitivity of outcome to the method chosen.

Secondary endpoints will not be sufficient to support efficacy in the absence of an effect for the primary endpoint. The protocol should propose a multiple testing strategy for secondary endpoints that adjust for multiplicity to be applied after the result for the primary endpoint is significant.

g. Statistical analysis plan

Before unblinding any phase 2b or phase 3 trial, sponsors should have in place a detailed finalized SAP. Although sponsors can update or modify an SAP as long as the trial remains blinded, sponsors should recognize that a detailed discussion may be needed concerning data access and appropriate firewalls for maintaining the integrity of the blind. If any major modification occurs, sponsors should discuss the modifications with the DAVP. Ideally, the SAP should be prepared at the time the protocol is made final, but we recognize that changes are sometimes made later, but before unblinding. The SAP should be considered as part of the protocol, and it can be either a section within the protocol (encouraged) or a separate document. The SAP should include the details on endpoint ordering, analysis population, structure of statistical hypotheses to be tested, statistical methods including the mathematical formulations, level of significance or alpha-level, alpha adjustments for multiple comparisons or interim analyses if applied, definition of visit window, handling of missing data, and sensitivity analyses.

It is important that the SAP prospectively identify the covariates to be used in the analysis. It is also important to choose covariates that are expected to strongly influence outcome.

Center-by-treatment interaction should be investigated and reported to assess consistency of the efficacy results.

[14] See the guidance for clinical trial sponsors *Establishment and Operation of Clinical Trial Data Monitoring Committees*.

h. Submission of data and programs

In the NDA submission, applicants should provide the complete or selected copies of original records that are usually portable document format files of the following:

- Case report forms (CRFs).
- Lab reports and randomization schedule.
- The standard operating procedure for randomization code generation.
- Screening dataset including the information on all patients screened.
- Raw datasets consisting of variables that come directly from CRFs or other original source documents.
- Analysis datasets including variables for key efficacy and safety analyses.
- Algorithms and programs used to create these analysis datasets directly from the raw datasets and programs for the primary and key secondary statistical analyses. If the analysis datasets were created from intermediate datasets other than original raw datasets from CRFs, applicants should provide the intermediate datasets and programs to cover both steps.

For additional information on regulatory submissions, see the guidance for industry *Providing Regulatory Submissions in Electronic Format — Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications*.

7. Accelerated Approval (Subpart H) Considerations

Traditional approval based on an endpoint of HIV-RNA suppression is the anticipated pathway for marketing approval. Suppression of HIV-RNA is a fully validated surrogate for HIV clinical disease progression. In addition, shorter term HIV-RNA changes are predictive of longer term HIV-RNA suppression in the setting of active antiretroviral drug regimens.

C. Other Considerations

1. Clinical Virology Considerations[15]

The clinical resistance analysis examines all virologic failure patients that experience viral rebound, have no antiviral response or an incomplete antiviral response, or discontinue before suppression. As such, the number of virologic failures in this analysis

[15] See the *Attachment to Guidance on Antiviral Product Development — Conducting and Submitting Virology Studies to the Agency: Guidance for Submitting HIV Resistance Data*.

may be different from the number of virologic failures in the snapshot approach analysis (see Appendix A). The examination of virologic failures in the clinical resistance analysis is designed to be more conservative to detect all possible signals and markers of resistance.

Proof-of-concept and efficacy trials should assess the development of HIV genotypic resistance to the investigational drug. Phenotypic and genotypic resistance testing should be performed on baseline and on-treatment failure samples (preferably the rebound confirmation sample) for patients who demonstrate virologic rebound (defined as a 1 $\log_{10}$ increase in HIV-RNA from nadir value or a confirmed HIV-RNA above 400 copies/mL after confirmed suppression to below 50 copies/mL). Any changes, including mixtures, in the amino acid coding sequence of the targeted genome region present in on-treatment or follow-up samples, but not in the baseline sample, should be reported as having emerged during therapy.

Genotypic resistance analyses should be performed on baseline samples from all patients in treatment-naïve and treatment-experienced trials to construct an effective background. In the case of new drugs from an established class, these data are important in evaluating the effect of transmitted or drug-selected baseline resistance-associated substitutions on response. In addition, baseline samples should be analyzed to identify HIV genetic polymorphisms that are associated with differential antiviral activity with the new drug. Phenotypic testing of a large subset of baseline samples also may be needed when an adequate genotypic resistance algorithm cannot be established.

Viral resistance-associated polymorphisms or substitutions observed in clinical trials but not identified and characterized in nonclinical virology experiments should be evaluated phenotypically by introducing the amino acid changes into the HIV genome, and determining the conferred fold-shift in susceptibility to the drug using appropriate cell culture and/or biochemical assays. In addition, phenotypic analyses of baseline and on-treatment failure clinical isolates should be analyzed and compared using a subset of trial patients representative of the HIV genetic diversity and virologic responses observed in clinical trials.

Sponsors should consider genotyping regions outside the direct HIV genome target depending on the characteristics of the antiviral drug and interactions of the target with other viral proteins. In cases when resistance is suspected based on viral RNA kinetics, but genotypic evidence of resistance is not detected, sponsors also should consider performing additional genotypic analyses using a method sufficiently sensitive to detect minority variants.

2. Pharmacokinetic/Pharmacodynamic Considerations

Trials conducted in HIV-infected patients should assess pharmacokinetics and the relationship between exposure and virologic suppression and toxicity in all patients.

Sponsors can use a combination of intensive and sparse sampling throughout development to characterize the pharmacokinetics of the investigational drug. For example, an intensive sampling schedule should be implemented in monotherapy trials. In longer term trials, however, an intensive sampling schedule might not be feasible, or may be feasible only in a subset of patients or over a limited period of time (i.e., a single assessment at steady state). Sparse PK samples should be obtained from as many patients in longer duration trials as possible, and the PK samples from these trials can be combined with intensive PK data from earlier trials for analysis. Sparse PK samples should be obtained at the time of virologic assessments, such as at weeks 4, 8, 12, 24, 36, or 48 or as otherwise specified in a protocol.

Sponsors can use the following two broad approaches to characterize the relationship between drug exposure and viral kinetics or virologic suppression of the investigational drug, depending on the development stage and purpose of the analysis. Both approaches allow for exploration of relevant covariates.

1. To aid the design of phase 2b or phase 3 trials, with respect to selection of dosage regimen, a mechanistic approach relating drug concentrations and viral kinetics is most appropriate. A mechanistic modeling approach should also account for the development of resistance to the investigational drug.

2. A simplified analysis relating proportion of patients with virologic suppression or virologic failure and appropriate exposure variable (e.g., minimum concentration or area under the plasma drug concentration versus time curve) can be used to support evidence of effectiveness and justify dose selection.[16]

Additional analyses of the exposure-safety relationship(s) using similar approaches as described in # 2 also should be performed to assist in evaluating the balance between effectiveness and toxicity of different dosage regimens.

3. Pediatric Populations

Under the Pediatric Research Equity Act (PREA), sponsors must study a drug in all relevant pediatric populations when submitting an application under section 505 of the Federal Food, Drug, and Cosmetic Act (FD&C Act) (21 U.S.C. 355) or section 351 of the Public Health Service Act (42 U.S.C. 282) for a new active ingredient, new indication, new dosage form, new dosing regimen, or new route of administration. However, the PREA requirements may be waived or deferred in certain circumstances.

Although a detailed discussion of how sponsors may comply with the PREA requirements is beyond the scope of this guidance, several points relevant to drugs for HIV treatment are addressed below. In addition, under the Best Pharmaceuticals for Children Act, drugs are eligible for 6 months of additional exclusivity if sponsors conduct pediatric clinical trials specified in a Written Request. New drugs for treatment of HIV

[16] See the guidance for industry *Exposure-Response Relationships — Study Design, Data Analysis, and Regulatory Applications*.

may be issued a Written Request if the FDA determines that information relating to the use of the drug may produce health benefits in relevant pediatric populations.

Early trials of antiretrovirals should enroll adult patients only, reserving drug administration to pediatric subjects until the pharmacokinetics, pharmacodynamics, and safety of the drug are reasonably well defined.[17] Sponsors are encouraged to begin discussions of their pediatric formulation and clinical development plan early in development, but pediatric clinical trials should be initiated after phase 2 adult data characterizing the safety profile and initial antiviral efficacy are available. To be in compliance with PREA, sponsors must submit a pediatric study plan to the FDA no later than 60 days after the end-of-phase 2 meeting.[18] If clinical trials in adults have demonstrated no significant safety concern that would preclude study in children, the pediatric development program should include, among other things:

- Development of an age-appropriate formulation.

- Clinical pharmacology trials to assess single- or multiple-dose pharmacokinetics (as appropriate for the drug) across the pediatric age range (2 weeks to younger than 18 years of age). Dose selection for the clinical pharmacology assessment and subsequent trials assessing efficacy and safety should be discussed with the review division.

- A sufficient number of patients in the pediatric safety database who have received the drug at the to-be-marketed dose or higher for at least 6 months to reasonably characterize the safety profile of the drug in pediatric patients. Generally, a safety database that includes 100 pediatric patients treated for at least 6 months will be sufficient but this number may vary based on drug-specific issues.

- A plan for long-term follow-up after treatment completion to assess growth and development, durability of virologic suppression. Follow-up over a period of at least 3 years is anticipated, but a postmarketing requirement provided after initial pediatric labeling also may be appropriate.

4. Early Access/Treatment INDs

Treatment INDs or other access protocols for antiretroviral drugs may be appropriate when sufficient clinical trial data have been generated to characterize a reasonably safe and active dose of an investigational drug. Ideally, the timing of a treatment IND is after phase 3 trials are fully enrolled or well underway so as not to interfere with phase 3 drug development. Treatment INDs can provide early access while phase 3 trials are being completed, analyzed, submitted, and reviewed by the FDA. Alternatively, individual

[17] See the guidance for industry *E11 Clinical Investigation of Medicinal Products in the Pediatric Population*.

[18] See section 505B(e) of the FD&C Act as amended by section 506 of the Food and Drug Administration Safety and Innovation Act of 2012.

patient INDs and treatment access protocols for intermediate size populations can occur earlier in drug development.

Historically, early access programs for the treatment of HIV infection allowed many patients to gain access to lifesaving drugs. However, for some individuals, early access to a drug amounted to sequential monotherapy and the emergence of multidrug resistance. Because treatment of HIV requires multiple drugs to achieve and maintain viral suppression below assay detection limits and to reduce the emergence of drug resistance to single drugs or drug classes, treatment INDs that include two or more investigational drugs or that allow co-enrollment in several treatment IND programs simultaneously are desirable. Treatment use of multiple investigational drugs should be supported by:

- Data and rationale that characterize the potential for PK-based drug interactions and potential for overlapping toxicity. Data to support dose modifications (if needed) when substantial drug interactions are present.

- Information suggesting the lack of antagonistic antiviral activity and minimal or no overlapping resistance profiles.

REFERENCES

Bartlett, JA, J Johnson, G Herrera et al., 2006, Long-Term Results of Initial Therapy with Abacavir and Lamivudine Combined with Efavirenz, Amprenavir/Ritonavir, or Stavudine, J Acquir Immune Defic Syndr, 43:284-92.

Centers for Disease Control and Prevention, 1992, 1993 Revised Classification System for HIV Infection and Expanded Surveillance Case Definition for AIDS Among Adolescents and Adults, MMWR, December 18, 1992, 41(RR-17).

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

Hill, AM, R DeMasi, and D Dawson, 1999, Meta-Analysis of Antiretroviral Effects on HIV-1 RNA, CD4 Cell Count and Progression to AIDS or Death, Antiviral Therapy, 3:139-145.

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

Marschner, IC, AC Collier, RW 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, JS, MR Elashoff, LC 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, FJ, KM Delaney, AC Moorman et al., 1998, Declining Morbidity and Mortality Among Patients with Advanced Human Immunodeficiency Virus Infection, N. Engl. J. Med., 338:853-60.

Saag, MS, P Cahn, F Raffi et al., 2004, Efficacy and Safety of Emtricitabine vs. Stavudine in Combination Therapy in Antiretroviral-Naïve Patients: A Randomized Trial, JAMA 292:180-9.

Strategies for Management of Antiretroviral Therapy (SMART) Study Group, WM El-Sadr, JD Lundgren, JD Neaton, F Gordin, D Abrams, RC Arduino, A Babiker, W Burman, N Clumeck, CJ Cohen, D Cohn, D Cooper, J Darbyshire, S Emery, G Fätkenheuer, B Gazzard, B Grund, J Hoy, K Klingman, M Losso, N Markowitz, J Neuhaus, A Phillips, C Rappoport, 2006, CD4+ Count–Guided Interruption of Antiretroviral Treatment, N Engl J Med, 355:2283–96.

Van Leth, F, P Phanuphak, and K Ruxrungtham, 2004, Comparison of First-Line Antiretroviral Therapy with Regimens Including Nevirapine, Efavirenz, or Both Drugs, Plus Stavudine and Lamivudine: A Randomized Open-Label Trial, the 2NN Study, Lancet, 363:1253-63.

.

APPENDIX A: RECOMMENDED APPROACH FOR EVALUATING VIROLOGIC RESPONSE IN CLINICAL TRIALS SUPPORTING ANTIRETROVIRAL APPROVALS

The time to loss of virologic response (TLOVR) method previously used in labeling by the DAVP for determining *virologic successes* at critical time points has often led to multiple queries between the DAVP and the applicant. Briefly, to be called a *virologic success* (HIV-RNA less than 50) by TLOVR,[19] a subject needed to have an HIV-RNA level below a detection limit on two time points and should not have experienced confirmed rebound (two time points) above the limit. This algorithm was, at times, cumbersome when subjects were less than perfectly adherent or when subjects needed to stop treatment for brief periods.

DAVP statistical and clinical reviewers recently completed a project titled "Handling uncertainty in endpoint selection and other endpoint issues." The goal of the project was to determine if simplified endpoints could be used for approval at Week 48. The team evaluated 18 trials from 7 NDAs with 8,046 patients. Results obtained using the TLOVR algorithm, which used data from every visit to consider the pattern of HIV responses, were compared to a less complicated *snapshot* approach that only used HIV-RNA data at the visit (window period) of interest. A high concordance between the TLOVR algorithm and snapshot results was observed. Using the TLOVR algorithm, 61 percent of the 8,046 patients remained in the study for 48 weeks and were virologic responders compared to 61 percent of the patients using the snapshot approach; 18 percent were virologic nonresponders using the TLOVR algorithm compared to 17 percent using the snapshot approach and approximately 20 percent discontinued before Week 48 using both approaches. Clinically significant differences between the two methodologies are minimal.

Based on the findings from the project and the ease of the snapshot method, pending supplemental NDAs and future NDAs should include virologic outcome results based on the snapshot approach in product labeling.

Snapshot Approach

For analysis of virologic outcome at a given time point, a window period for possible virologic assessments can be used as follows:

- Window size is ½ the duration of time between study visits.

- Windows can be smaller at earlier time points than later time points.

[19] Previously, labels used the term *virologic success* or *virologic failure* to describe subjects who had HIV-RNA levels below or greater than or equal to 50 copies, respectively. However, we now prefer not to use the terms success or failure, but rather just state whether the viral load was below or greater than 50 copies. Transient blips of HIV-RNA greater than 50 copies occur for a variety of reasons and this does not always signify true virologic failure to the regimen. True virologic failure may only be determined after assessment of drug adherence, repeat HIV-RNA testing with continued treatment, and/or resistance testing. Snapshot time windows allow time for clinical assessment and retesting to reduce the number counted as greater than 50 copies because of transient blips.

- If trial-defined windows differ from the proposed windows in Table A, alternatives should be discussed with the DAVP. In most cases the protocol-defined windows for completed trials are acceptable; however, for future trials we encourage standardization and recommend the windows in Table A.

Table A: Proposed Windows

Visit	Window (Through End of Study Week) *(Express in Days for Nonoverlap)*	Window (Days)
24	18-30	126-209
48	42-54	294-377
96	90-102	630-713

Table B is an example of efficacy presentation in labeling.

Table B: Virologic Outcome at 48-Week Window (294 to 377 Days)

	Drug A	Drug B
HIV-RNA < 50 copies/mL[±]	60%	50%
HIV-RNA ≥ 50 copies/mL[#]	20%	30%
No Virologic Data at Week 48 Window		
Reasons		
Discontinued study/study drug due to AE	10%	8%
or Death[*]	6%	6%
Discontinued study/study drug for Other	4%	6%
Reasons[**]		
On study but missing data in window		

[±] **Assays with other lower limits also can be used.**

[#] Includes patients who changed any component of background therapy to a new drug class or changed background components that were not permitted per protocol or changed any background drug in the regimen because of lack of efficacy (perceived or documented) before Week 48, patients who discontinued study drug or study before Week 48 for lack or loss of efficacy and patients who are equal to or above 50 copies/mL in the 48 week window

[*] Includes patients who discontinued because of adverse event (AE) or death at any time point from Day 1 through the time window if this resulted in no virologic data on treatment during the specified window.

[**] Other includes: withdrew consent, loss to follow-up, moved, among others.

Principles of snapshot analysis

Some general concepts of the snapshot approach include the following:

- The primary efficacy endpoint should be primarily a virologic endpoint and not a clinical endpoint. This method follows a *Virology First* hierarchy.

- Because this is primarily a virologic endpoint, the hierarchy for assessing row and column percentages is HIV-RNA below 50 copies/mL or HIV-RNA greater than or equal to 50 copies mL, first, for any given time window followed by reasons for *No Virologic Data in the 48-Week Window*.

- Percentages not included in the HIV-RNA below or greater than or equal to 50 copies/mL rows should describe reasons for no data at a specified analysis time window in the AR population. These percentages should not represent comprehensive safety or clinical efficacy analyses.

Procedures for calculating virologic outcome

The following examples use a detection limit of 50 copies/mL, but approved sensitive assays with other detection limits also can be used.

- **Data in the window**

  Virologic outcome should be determined by the last available measurement while the patient is on treatment and continued on trial within the time window (see Table A).

  – Examples: HIV-RNA = 580 copies/mL at Day 336, HIV-RNA below 50 copies/mL on Day 350. This should be categorized as HIV-RNA below 50 copies/mL.

  – In the rare example that someone would have HIV-RNA below 50 copies/mL at Day 336 and then equal to or above 50 copies/mL at Day 350, this would be considered a failure (we believe this will be rare, because undetectable patients would not likely have a second lab result in a window).

- **No data in the window**

  – If there are no data in a time window, then percentages for each category of missing data should be tallied.

  – There are three reasons for no data in the window:

    1. **Discontinued study due to Adverse Event or Death.** Any patient who discontinues because of an AE or death **before** the window should be classified as *Discontinued due to AE* or *Death* (as appropriate), regardless of the HIV-RNA result, even if the HIV-RNA is below 50 copies/mL at the time of discontinuation.[20] However, if a patient has an HIV-RNA value in the time window and also discontinues **in the time window**, the viral load data should be used to classify the patient's response. This is the Virology First hierarchy. Example: HIV-RNA below 50 copies/mL at Day 336 and discontinues because of AE or even dies on Day 360 — this person is categorized as having HIV-RNA below 50 copies/mL. Likewise if HIV-RNA is 552 copies/mL on Day 336 and the patient discontinues on Day 360, the patient is categorized as having HIV-RNA greater than or equal to 50 copies/mL.

[20] There should not be a separate category for Death. We believe a separate category for Death is misleading, because it does not account for all deaths in the trial. Instead, text describing percentages of deaths can be included in the CLINICAL STUDIES section of product labeling.

2. **Discontinued study for Other Reasons.** The examples above also apply to this category. If a patient discontinues the study before the window because of *lack of efficacy* then the patient should be included in the HIV-RNA greater than or equal to 50 row and not in the *Discontinued for Other Reasons* row. To further clarify, for patients who Discontinued for Other Reasons, it is important to realize that in the Virology First hierarchy only patients who have achieved virologic suppression can be counted as Discontinued for Other Reasons. If a patient discontinues because of *subject withdrew consent* and his or her HIV-1 RNA result at the time of discontinuation was equal to or above 50 copies/mL, then he or she should be categorized as HIV-RNA greater than or equal to 50 and NOT as Discontinued for Other Reasons. However, if a patient discontinued because of *Lost to Follow-Up* and the last HIV-RNA result was 49 copies/mL, then the patient can be categorized as Discontinued for Other Reasons.

   Likewise, if patients changed background treatment — *not permitted by protocol* — they should be considered an efficacy failure and captured in the HIV-RNA greater than or equal to 50 copies/mL row.

3. **On study but missing data in window.** Only data in the window can be used for patients remaining on study. For example, if there are no data during Days 294 to 377, but there is an HIV-RNA below 50 copies/mL on Day 380, this patient should be considered *On Study but Missing Data in Window*. This patient can count as below 50 copies at subsequent analysis points (e.g., 96 weeks), if he or she remains undetectable at the subsequent analysis window (e.g., 96 weeks). Conversely, if there are no data during Days 294 to 377, but there is an HIV-RNA equal to or above 50 copies/mL on Day 280, this patient also should be classified as On Study but Missing Data in Window.

Optimized Background Therapy Substitutions After Randomization

Typically trials have permitted one in-class substitution of an optimized background therapy (OBT) drug for documented toxicity reasons. As more drugs became available, cross-class substitutions were permitted in some trials; however, drug substitutions potentially can affect long-term durability of a regimen particularly if the OBT change occurred later in the trial. OBT substitutions (in-class or cross-class) permitted per protocol for documented toxicity reasons can be permitted on or before the first trial visit without penalty. If OBT substitutions for toxicity reasons occur after the first trial visit, then patients should be categorized as having HIV-RNA greater than or equal to 50 copies/mL if they have HIV-RNA above 50 copies/mL at the time of switch.

Applicants have asked to amend the algorithm such that only cross-class switches are classified as primary endpoint failures because not allowing in-class OBT substitutions may create disincentives. Specifically, investigators may not have incentive to ensure follow-up after an OBT switch because those patients are deemed as analysis failures, or investigators may unnecessarily increase early switches to avoid classifying patients as failures in the primary efficacy analysis.

We decided not to amend the algorithm for the following reasons:

- All in-class switches are not the same. With the expanded number of drugs in each class and the approval of second generation drugs within the same class, switching therapy after knowledge of viral load changes may confound the results. One would then have to decide which switches are appropriate for the population being studied.

- We attempted to make the snapshot as concise and stringent as possible to reduce the amount of end-of-FDA-review negotiations over single cases. Having to decide which in-class switches are appropriate for specific populations (e.g., naïve, experienced) would complicate the algorithm. Example: In what population is a switch from atazanavir to darunavir considered acceptable?

- We believe that the unwanted scenarios mentioned above can be minimized. Both types of analyses can be performed, perhaps allowing cross-class switches in sensitivity analyses. However, for FDA labeling purposes, the snapshot should be used. Therefore, investigators could be informed that not all analyses may result in their particular patient counting as a *failure* if he or she switches background drugs and that follow-up should be maintained.

- We do not believe that there is one *correct* analysis. All analyses only approximate truth. The snapshot approach strives for efficiency and consistency across multiple applications. This should not prohibit academic investigators from presenting a variety of analyses at scientific meetings. Differences can be described.

Datasets for Snapshot Approach

For a submission with multiple trials, each trial should have its own dataset for the snapshot analysis. The datasets should contain, at minimum, the following information:

- Study identification (ID)

- Patient study ID

- Study day and date of last double-blind treatment

- Virologic outcome based on the snapshot approach (i.e., HIV-RNA below 50 copies/mL, HIV-RNA greater than or equal to 50 copies/mL, discontinued due to AE or death, discontinued for other reasons, on study but missing data during window)

- The HIV-RNA measurement and the corresponding study day and date used to determine the above virologic outcome if the measurement was not missing

- Study day and date when the patient switched to open-label treatment because of lack or loss of virologic suppression, if applicable

- Discontinuation study day and date, reason for discontinuation, and last on double-blind, treatment measurement before discontinuation for the patients who discontinued drug

The treatment phase in the dataset should be defined and only include three categories as follows: screening (or baseline), treatment, and follow-up.

APPENDIX B: NONINFERIORITY MARGIN JUSTIFICATIONS

1.0 Justification for a Noninferiority Margin Using EFV as a Control Arm in Treatment-Naïve Studies on a Background of Dual Nucleoside Therapy

The noninferiority margin for comparing the *potent anchor drug* or *third drug* in regimens for HIV treatment-naïve patients is 10 to 12 percent. This margin is an M_2 delta, based on the treatment effect we clinically wish to preserve compared to active controls. We have known for years, based on well-controlled superiority trials, that an M_1 for assessing comparability to a PI or NNRTI as a third drug added to a dual nucleo(t)side background is large (approximately 45 percent — using lower confidence bounds for the endpoint of HIV-RNA below 50 or 400 copies/mL at 48 weeks). The rationale is as follows.

1.1 EFV's treatment effect is highly reproducible and dual nucleosides alone are known to be suboptimal for durable virologic suppression

Few individuals (approximately 2 percent or less) receiving only two nucleoside analogues achieve viral load suppression below a 400 copies/mL detection limit. Even fewer suppress HIV-RNA below 50 copies/mL. The few that suppress below the detection limit are those individuals with low baseline viral loads below 5,000 copies and high CD4 cell counts. These people are known as long-term nonprogressors but few enroll in registration trials. Beginning in 1995, suppressing viral load below assay detection limits was a new phenomenon, recognized when PIs and NNRTIs became available and were added to a dual nucleo(t)side backbone. Before PIs and NNRTIs, long-term suppression (less than 24 to 48 weeks) of viral load was virtually unheard of. The addition of a PI or an NNRTI to two nucleosides basically converted a negligible viral load response (less than 2 percent) to a response rate of 60 to 90 percent, owing to the potency of PIs and NNRTIs, marked antiretroviral synergy of an antiviral regimen, and a formidable resistance barrier that three drugs confer compared to two drugs.

Several current drug labels contain examples of response rates observed with dual nucleoside therapy. All of these studies show that dual nucleoside therapy is associated with a negligible response rate (defined as suppressing viral load below an assay limit). The genetic barrier for two nucleo(t)side analogue drugs is known to be insufficient to durably suppress viral load in most individuals based on calculations of reservoirs, replication rates, and potential for pre-existence of antiretroviral mutations. Examples of dual nucleoside response rates are listed in Table C.

Table C: Virologic Response Rates for Dual Nucleoside Studies (Approximately 48 Weeks)

Drug Label Study	Nucleoside Backbone	Nucleoside Response Rate < 400 at 48 Weeks	Triple Response Rate
Nelfinavir -Study 511	ZDV/3TC	3%	58%
Indinavir -ACTG* Trial 320	ZDV/3TC	2%	45%
Indinavir Merck Trial-035	ZDV/3TC	0%	80%

* AIDS Clinical Trial Group

EFV has been extensively studied in triple regimens in clinical studies of 48 weeks duration in treatment-naïve patients and was part of the control regimen in many of these studies. In Table D, response rates for proportion below 400 copies/mL for triple regimens that included EFV ranged from 64 percent to 84 percent, and for proportion below 50 copies/mL ranged from 37 percent to 80 percent. (Note that the 37 percent response rate is an outlier and samples were believed to be mishandled in that study; without this study the range is 59 to 80 percent). There has never been a study in treatment-naïve individuals in which EFV and two nucleosides did not perform in this range. In contrast, dual nucleo(t)side treatment consistently showed a response rate of less than 5 percent. Therefore, the treatment effect for EFV is reliably around 60 to 80 percent and with the use of fixed-dose combinations has been closer to 80 percent.

Table D: Virologic Response Rates for EFV-Based Regimens

Drug Label (or Reference) Trial	Regimens	Response Rate < 400 (50) Copies/mL at 48 Weeks
(Bartlett et al. 2006) CLASS Trial	ABC/3TC/EFV ABC/3TC + AMP/ritonavir ABC/3TC + d4T	81% (72%) 75% (59%) 80% (60%)
Atazanavir Study AI 424-034	ZDV/3TC + ATV ZDV/3TC + EFV	70% (32%) 64% (37%)
Efavirenz Study 006	ZDV/3TC + EFV ZDV/3TC + IDV IDV + EFV	70% (64%) 48% (43%) 53% (47%)
(Van Leth et al. 2004) 2NN Trial	D4T + 3TC + NVP d4T + 3TC + NVP d4T + 3TC + EFV d4T + 3TC + EFV + NVP	(70%) (65%) (70%)
Abacavir CNA 30024	ZDV/3TC + EFV ABC/3TC + EFV	71% (69%) 74% (70%)
(Saag et al. 2004) Study 301A	FTC +ddI + EFV D4T + ddI + EFV	81% (78%) 68% (59%)

continued

Table D, continued

Drug Label (or Reference) Trial	Regimens	Response Rate < 400 (50) Copies/mL at 48 Weeks
Tenofovir Study 903	TDF + 3TC + EFV D4T + 3TC + EFV	80% (76%) 84% (80%)
Tenofovir Study 934	TDF + FTC + EFV ZDV + 3TC + EFV	81% (77%) 70% (68%)
Lamivudine EPV20001	ZDV+ 3TC (bid) + EFV ZDV + 3TC (qd) + EFV	65% (63%) 67% (61%)
Abacavir CNA 30021 Study	ABC (bid)+ 3TC + EFV ABC (qd) + 3TC + EFV	(68%) (66%)

One should note that by 48 weeks the proportion below 50 copies/mL and proportion below 400 copies/mL are fairly similar for most EFV regimens, within 10 percent and usually within 5 percent, except for one outlier mentioned above.

In the trials above, the dual nucleo(t)sides ABC+3TC, d4T+3TC, TDF+3TC (or FTC), and ZDV+3TC with added EFV, performed similarly. TDF+FTC has on occasion performed slightly better, but in some cases treatment effect may be driven by better tolerability rather than virologic response.

1.2 EFV has been shown to be superior to two older PIs that are well known to be active controls responsible for the sharp decline in AIDS mortality in the last decades.

In previous studies two nucleosides plus indinavir (IDV) has been shown to be superior to two nucleosides alone at approximately 48 weeks (proportion below 400 copies/mL). In ACTG 320, ZDV+3TC+IDV was superior to ZDV+3TC by approximately 40 percent. In the Merck study 035, ZDV+3TC+IDV was superior to ZDV+3TC by 80 percent (+/- 18 percent);[21] therefore, the lower confidence bound is 62 percent. In Study 006, EFV was superior to the known active control IDV by 21 percent (+/- 11.5 percent) for proportion of patients achieving below 50 copies/mL. Therefore, the 95 percent lower confidence bound for EFV compared to a highly active control is 10.5 percent. Therefore, the contribution of EFV is probably at least 10 percent more than the treatment effect of IDV.

We are recommending a noninferiority margin (M_2) of 10 to 12 percent, which is much less than the lower bound of the treatment effect of either EFV or IDV based on historical studies. An M_2 of 10 to 12 percent is clinically reasonable because it preserves a large portion of the treatment effect. In addition, in the setting of ongoing monitoring of viral load, failing therapy may be detected sufficiently early to allow individuals to change their regimen and avoid clinical consequences of disease progression.

[21] 1.96 times the standard error of the risk difference

Other support for EFV comes from studies in which EFV was superior to nelfinavir (NFV) in both a treatment-naïve (ACTG 384) and treatment-experienced study. NFV is known to be superior to ZDV+3TC by a margin of 55 percent (+/- 2 percent); lower bound 53 percent.

2.0 Justification for a Noninferiority Margin Using an NRTI as a Control Arm in Treatment-Naïve Studies

As stated in section III.B.3., Choice of Controls, investigational NRTIs should be compared only to control NRTIs in the context of an NNRTI-based regimen. Because boosted PIs have a high genetic barrier to resistance and a substantial proportion of patients may achieve undetectable HIV-RNA levels with a boosted PI alone, the quantitative contribution of an NRTI to a boosted PI regimen is unknown. Likewise, the quantitative contribution of an NRTI to an integrase strand transfer inhibitor-based regimen is also unknown because of limited numbers of studies with this drug class. First generation NNRTIs, however, are known to have a low genetic barrier to resistance and when used as monotherapy, nearly 100 percent of individuals will develop resistance in a matter of days to weeks. This has been documented for nevirapine, and based on a similar resistance profile is believed to be the same for EFV. Therefore, because of synergy, nearly all of the response rate in an NNRTI-based regimen also can be attributed to the two nucleo(t)side components of the regimen.

Based on early studies with NNRTIs such as nevirapine and delavirdine, one NRTI in combination with an NNRTI was not sufficient to achieve and maintain undetectable HIV-RNA levels. Conservatively one could attribute half of the treatment effect to each NRTI. In two recent trials in treatment-naïve patients, the lower bound for the treatment effect for an EFV/tenofovir/emtricitabine regimen was 77 percent (pooled data from two trials). Therefore, half of the treatment effect (38 percent) could be attributed to each NRTI. If one wanted to preserve an additional 50 percent of the effect, the margin is 19 percent. However, clinically we do not want to lose more than 10 to 12 percent of the treatment effect (M_2 margin). Similarly, for the reasons stated, an M_2 of 10 to 12 percent is an acceptable margin for an endpoint of HIV-RNA below 50 copies/mL at 48 weeks.

3.0 Justification for Noninferiority Margin in Treatment-Experienced Studies

The justification of a valid noninferiority margin in treatment-experienced trials is based on past performance of the active control and comparison of prior trial conditions to the current trial. The noninferiority margin determination for HIV treatment-experienced trials is complicated by variations in response rates across trials, use of different background drugs, and differences in baseline patient characteristics. The noninferiority margin should take these variables into account and a new protocol should attempt to replicate the original superiority trial for the active-controlled drug with respect to patient characteristics and protocol procedures. One issue encountered in establishing a noninferiority margin includes the change in virologic response rates for optimized background regimens over time. As presented in Table E, the proportion of patients with HIV-RNA below 50 copies/mL from the optimized treatment regimen (control) in three recent trials to support approval of these new drugs increased from 2004 to 2008. As

expected, the patient characteristics, namely the phenotypic susceptible score (PSS) at baseline,[22] influenced the response rates.

Table E: Virologic Response (HIV-RNA Below 50 Copies/mL) for OBT (Control) Over Trials/Time

Drug/Trial/Time	PSS=0	PSS=1	PSS=2	PSS ≥ 3
Maraviroc Motivate Trials 2004-2006	3%	5%	7%	42%
Raltegravir Benchmark Trials 2006-2007	2%	29%	39%	61%
Etravirine DUET Trials 2005-2008	6%	32%	62%	75%

Sponsors are encouraged to provide detailed supporting documentation for noninferiority treatment-experienced trials early in the protocol development stage. The proposed noninferiority margin should be discussed with the FDA at the time of submission of the protocol for FDA comments.

[22] A PSS is the number of drugs to which a patient's virus is susceptible according to phenotypic laboratory resistance tests. A score of zero means that the patient has no remaining drugs to which his or her virus has full susceptibility.