The Challenges of Clinical Trial Design in Assessing the Effects of Anti-HIV Therapy in Heavily Pre-treated Patients

Table of Contents

The Workshop organizers would like to thank several people for their assistance, input and guidance. The Scientific Co-Chairs of the Workshop Planning Committee were Roy Gulick, MD and Doug Richman, MD. John Mellors, MD and Julio Montaner, MD also provided significant support. We would also like to thank: Janet Anderson, MD, Terry Blaschke, MD, Jeff Chodakewitz, MD, Victor De Gruttola, MD, Heidi Jolson, MD, Jim Neaton, Ph.D., and Carla Pettinelli, MD for their assistance in the development of the panel presentations. Bio-Chem Pharma, Gilead Pharmaceuticals, Glaxo Wellcome, and Pharmacia & Upjohn provided support to Project Inform and TAG for the conference.

David Barr – Forum for Collaborative HIV Research
William Duncan, Ph.D. – NIAID Division of AIDS
Ben Cheng – Project Inform
Spencer Cox – Treatment Action Group

The Challenges of Clinical Trial Design in Assessing the Effects of Anti-HIV Therapy in Heavily Pre-treated Patients

The goal of this meeting is to discuss issues in the design and implementation of studies of salvage therapy regimens in heavily pre-treated patients. The meeting will allow for the presentation of differing needs, priorities and challenges faced by industry, researchers, regulators and patients conducting and participating in this research. At the meeting, we will try to agree upon some definitions for treatment failure and success. Further, we will work to develop a better understanding and agreement between parties regarding what is necessary and feasible when designing studies of new drugs for indication in salvage therapy.

EXECUTIVE SUMMARY

On May 25th, 65 people met to discuss the design of studies that evaluate therapeutic options for people with HIV in whom current anti-retroviral therapy is no longer working. The group consisted of leading researchers from the U.S., Canada, and Europe, drug regulators, administrators, patient advocates, health care providers, and representatives from the pharmaceutical industry. The meeting was sponsored by the Forum for Collaborative HIV Research, Project Inform, Treatment Action Group, and the NIAID Division of AIDS. The Scientific Co-Chairs of the meeting were Roy Gulick, M.D. of Cornell University Medical Center and Doug Richman, M.D. of the University of California/San Diego.

For all the participants, data on the rising rates of virological failure comes as no surprise. The numbers of patients for whom antiretroviral therapy is failing is growing steadily. There is an urgent need to address the treatment options for those patients now and in the future. This meeting sought to catalyze action among the relevant parties to move forward efforts in research and drug development.

The results of the meeting highlight the complexity of the problem. Participants did an excellent job of identifying the relevant research questions and the obstacles – both scientific and logistical - we face in getting answers to those questions. The meeting made clear that overcoming those obstacles is a great challenge. Many approaches were suggested, though we reached no overriding consensus on how to move forward. It is our hope that the meeting provided an opportunity for all participants to better understand the

various issues involved in studying these questions, and gave people ideas for further research. Future discussions need to clarify and expand upon these ideas. A recent follow-up meeting addressing strategies for treatment interruption is a good example of such discussion.

While there remain many key unanswered questions about how to address failure of the primary regimen, e.g. when to switch, there was general agreement that patients failing their first regimen would look to a potent regimen of new drugs, with at least one drug from a previously unused class. In two studies, genotyping has been shown to be helpful in choosing the new regimen. The focus of our discussion was patients who have already been through several antiretroviral treatment regimens. There is little data to guide physicians and patients in their treatment decisions. In order to better understand how to best treat these patients, two central questions emerge: (1) how to do the best with the drugs we have, and (2) how to incorporate new drugs.

In assessing the value of new drugs for salvage regimens, an overarching problem is the competing need to design studies geared towards drug approval and studies designed to improve the clinical management of patients. Registrational studies are generally based on 24 weeks of virological data. Clinical management studies require larger numbers of patients, followed for longer periods of time with clinical disease as well as virological and immunological endpoints. Both kinds of studies are important. Long-term data on clinical outcomes are needed to understand the consequences of multiple changes of treatment and the timing of change. Short-term data assessing the effect of adding a new drug is one way to assess in evaluating the activity of a new drug, but may be dangerous as it may lead rapidly to resistance.

Companies developing new products have limited supplies of study drugs. Decisions about which studies to undertake are complicated, with priority often given to those studies that can bring a new drug to market quickly. Specific labeling indications for salvage therapy use would provide incentive to industry to design salvage studies prior to approval. The FDA can assist by providing direction to industry regarding standards for

approval specifically of salvage therapy labeling. When an endpoint of undetectable viral load is not feasible, as would often be case for salvage therapy indications, changes in viral load could be used, but such changes would probably need to be for sustained longer than 24 weeks in order to suggest clinical significance.

Participants agreed that studying multiple experimental agents in salvage therapy regimens is important if the risks of adding single agents is to be avoided. Several factors confound the ability to do such studies including:

- the absence of an accepted standard of care for heavily pre-treated patients;
- the current FDA standard of distinguishing the relative contribution of each agent presents vexing trial design issues for companies developing new products;
- drugs are made by multiple companies requiring inter-company collaboration in drug development plans;
- the need for more pharmacokinetic and interaction studies at an earlier stage of drug development;
- the different rates at which manufacturers scale-up product supply;
- the difficulty of achieving consensus on study protocol designs;
- the need for rigorous confidentiality agreements.

New drugs could be first studied using multi-staged accrual designs where single drugs are tested for 1 –2 weeks in small numbers of patients for virological effect. Those drugs that show the desired level of activity could then be used in longer-term and/or larger studies. However, in order to assess the pharmacokinetics of new drugs in such a study, it can take almost two weeks to reach a steady state needed for study. Screening studies to test drug activity could be nested into longer-term management trials.

The definition of treatment failure is obviously a key question for study design. A plasma HIV RNA level above the level of level of detection may be an appropriate definition for first-line failure, but the criteria used to decide when to change treatments may differ from first to second to third-line failure. Drugs may be providing a therapeutic effect in patients who are unable to achieve undetectable plasma HIV RNA levels. Studies to address when to switch therapy could compare changing vs. not changing, changing at different RNA levels, or at different CD4 levels.

The design of regimens to which patients should be switched is complicated by many factors including the varied treatment history of patients, multiple drug-drug interactions, overlapping toxicities, and confusion about how to interpret results from resistance assays. Understanding the cause of treatment failure is particularly difficult and must be assessed for patients individually. For example, was failure caused by pre-existing resistance mutations due to sub-optimal treatment, poor adherence, or inadequate absorption? A classification system to reduce patient variability by using blocking factors to compare more similar patients was recommended.

There was significant discussion about studying three different treatment options if it is not feasible to switch to a new three-drug regimen: (1) the value of adding a new drug or drugs to a current regimen that may still have anti-viral activity, (2) the value of continuing a current regimen despite evidence of virological failure, and (3) interrupting treatment until enough new drugs are available. Given the large number of possible combination regimens to be assessed, factorial and partial factorial trial designs may be particularly useful.

The use of resistance testing, particularly to define patient inclusion in clinical studies, raises a series of important but vexing questions, including:

- How do you measure resistance in previously treated patients who have been off therapy for a significant period of time?
- The lack of an established system for classifying drug resistance;
- Real time testing is slow and expensive, especially in multi-site studies.
- Exclusion of people on the basis of specific mutations or phenotypic results limits the ultimate generalizability of the findings of the study.
- Very specific mutation criteria requires more screening, making studies costly and slowing accrual.
- We do not understand the prognostic importance of different mutations.
- It might be useful to have multiple studies – one for people who meet specific geno- or phenotypic entry criteria, and others for those ineligible for that study.

Salvage studies of specific drugs or regimens are best for short-term study with virologic endpoints because an expected large group of patients who would fail the new regimen virologically could not be kept on a specific regimen after failure occurs. To carry out

large, long-term strategy studies in a timely manner, the comparison of classes of drugs may be more feasible and appropriate.

The lack of and need for drug interaction data was cited as a particularly pressing problem. As more drugs become available, the need for drug interaction data grows. The problem is compounded by the use of drugs for opportunistic infections, non-HIV related treatments, and other drugs. One recommendation was the development of a data collection system within clinical trial networks to gather data on demographics, creatinine clearance, and hepatic function, and to collect blood samples to measure drug levels which can be sent to a central laboratory and analyzed so that there are more informative data on drug interactions. Workshop participants recommended the formation of a consensus group to focus on standardization of methodologies for blood collections, validation of assays, and issues around drug interaction when steady state levels are achieved. Another recommendation is to have a clearinghouse for drug interaction data. Pharmaceutical companies should collaborate to get drug interaction data earlier in the drug development process.

Therapeutic drug monitoring can be useful for several purposes including the measurement of adherence, to vary the intensity of drug exposure in relation to the IC90 value, in order to optimize antiviral effects and/or minimize toxicities.

Finally, several conference participants worked together to produce a study proposal to evaluate several experimental agents simultaneously, while offering study participants a reasonable chance of therapeutic success. The rough outline is this:

A) SOC + Drug X + Drug Y + Drug Z
B) SOC + Drug X + Drug Y
C) SOC + Drug X + Drug Z
D) SOC + Drug Y + Drug Z
E) Drug X + Drug Y + Drug Z

MEETING SUMMARY

Opening Discussion

The workshop opened with a wide-ranging discussion of the issues that arise during the evaluation of new therapies in heavily pre-treated patients. For instance, as one workshop participant put it, when looking at available salvage therapies, "standard of care sucks." Carlton Hogan, who provided the patient's perspective in the opening session, first outlined the difficulty of defining a patient who is experiencing treatment failure. Is it someone resistant to a single protease inhibitor, a class of drugs, or multiple classes of drugs? Does it include someone initially infected with multi-drug resistant virus? He suggested that for the purposes of this discussion, we focus on patients with fewer treatment options. These patients face multiple obstacles in obtaining effective anti-viral therapy. There are fewer clinical studies for them to enter. Because they are often in later stages of HIV disease, they need to take several different drugs for prophylaxis, in addition to antiretroviral therapy, and often suffer from more adverse reactions to therapy. These patients also need to be particularly concerned about the effects that any change in therapy may have on their treatment options for the future.

The poor efficacy and tolerability of current salvage regimens means, in turn, that there is no widely accepted control arm to which new regimens could be compared. In addition, it creates difficulty in enrolling patients, who are often unwilling to enroll in studies where treatment may involve an unacceptably low likelihood of success and the potential for increased multi-drug resistance. Clinicians may also be unwilling to recommend clinical trials to patients when treatment success is unlikely and the risks of adverse effects are high.

Additionally, some participants suggested the current regulatory structure inadvertently creates incentives to avoid assessment of new therapies in heavily pre-treated patients. The Food and Drug Administration does not require such studies, and many companies perceive that a negative finding could adversely affect their chances for rapid marketing approval. A representative of FDA pointed out that the agency is strictly governed by

its authorizing legislation, and has very limited power to compel manufacturers to conduct salvage therapy studies. While the agency certainly could attempt to limit labeling language so that it more narrowly reflects the studies presented in the approval package, such limited indications often present subsequent concerns about third-party reimbursement. FDA does, however, strongly encourage companies to evaluate new therapies in several different patient populations, including heavily pre-treated patients, and this representative expressed a willingness to work with manufacturers to ensure that such data contribute to rather than hinder their overall product development efforts.

Dr. John Mellors made a brief presentation in which he proposed rapid, short-term evaluation of virologic potency of large numbers of therapeutic combinations. Such studies could, Dr. Mellors noted, look at two- or four-week virologic response, and could attempt to correlate that response to baseline genotype and phenotype, as well as offering opportunities for assessment of pharmacokinetic interactions. More successful combinations could then be compared in larger-scale controlled clinical trials. Workshop participants were enthusiastic about this proposal, and returned to the idea several times during the ensuing panel discussions.

One participant commented that there are approximately 20 new agents on the market that can be used for treatment. This presents the potential for a "Latin square" study design, whereby several multiple new agents that can be used in combination with one another. This participant expressed the opinion that the HIV community can not afford to wait for the development of adverse events, but must obtain the regulatory freedom to utilize various multiple drug combinations based on pharmacokinetic data.

Another participant pointed out that there are two goals associated with drug development: 1) new drugs must be shown to be active, safe and effective; and 2) optimal treatment regimens need to be defined. The participant then posed the question, "How do we obtain both of these goals without them compromising each other?" The participant went on to comment that parallel studies, some assessing different regimens in pre-

treated populations and other assessing the safety and efficacy of specific new drugs, should be encouraged. While some companies perform parallel studies, most do not.

Another participant expressed the concern that new regulatory requirements for salvage assessment will mean longer approval periods for new drugs. In response, a workshop attendee suggested that perhaps multi-factorial pivotal trials with two companies should be required during phase III development. If both drugs appear to be very active, this is certainly interpretable; however, if either drug were inactive, results would be "less than normal". Determining which drug was most active would require further research. How the participant would define “very active” and “less than normal” was unclear. A participant asked if each of these studies could conceivably have independent classification as a clinical trial. A regulator responded that this would be quite conceivable, as long as the studies were proposed in this manner.

Changing subjects, one participant expressed the opinion that perhaps the issue was being confused, and the HIV community and physicians as a whole should begin thinking about advancing their knowledge of new drug regimens, irrespective of regulatory approval.

One industry representative commented that "industry is struggling with many different factors, and we must look at all of these factors." This participant felt that genotypic and phenotypic testing should enhance salvage therapy studies. Naive patients cannot be “written off” in the development of new drugs; often 60% of the patients do not achieve durable responses. Industry is in need of a more innovative approach with new technology.

Another industry representative commented that industry commit themselves to combining multiple drug regimens in trials, based on pharmacokinetic data. The industry representative concluded the session by saying that the current concept of re-using drugs in previously treated patients is easy for short periods of time (such as four to eight weeks) if a temporary drop in viral load is the desired outcome. But, if we are looking for a long-term solution toward better patient outcomes, it is essential that we "choose the

winners, ditch the losers, and take advantage of what we can learn about adherence and toxicity."

Statistical Considerations in Trial Design

The goal of this panel was to raise issues regarding the difficulties in clinical trial design to test the effectiveness of anti-retroviral therapy in patients for whom current regimens are failing and to present possible design strategies for discussion.

Janet Darbyshire, O.B.E.

Dr. Darbyshire outlined two basic needs: the need to learn about how to do the best we can with the drugs we have, and the need to learn how to incorporate new drugs. The latter requires an examination of both how to improve therapy for people and an assessment of new products.

The first challenge is to define what we mean by treatment failure – virological, immunological, and/ or clinical. Are we discussing failure of the first regimen or failure of the second or third line? There is little comparative data about first line failure; there is even less regarding subsequent failures. Despite some data pointing to a better outcome if therapy is changed quickly after evidence of virological failure, the benefits of early vs. late change are still unknown, especially over the long-term. The criteria one uses to decide when to change therapy may differ from the failure of a first regimen to the failure of a second, third or fourth regimen. Long-term data on clinical outcomes are required to evaluate when to change therapy. Short-term data regarding the effect of adding a new therapy may be important for assessing the activity of the new drug, but carries the risk of rapid emergence of resistance to that drug if it is added to a "failing regimen".

There is still no general agreement about when to change a failing first line regimen, but decisions about when to change subsequent regimens are even more difficult. Studies to address this question could compare changing vs. not changing, changing at different

RNA levels, or at different CD4 cell levels. The important thing is that there needs to be enough difference between the two courses of action in order for a difference to be detected. Testing different times to change or different definitions of failure requiring a treatment change is an important possible area of study.

The next set of questions regard what treatments to switch to. Do you switch all therapies, if possible? Do you add drugs to the current regimen? If the current drugs are still having some effect, adding one or two new drugs to a regimen that is still having an impact on viral fitness may be a good thing to do. If we intensify treatment by including the use of new drugs, should we do so early in the failure of a regimen or wait until "full blown failure"?

What about stopping? Are there any benefits from a temporary stop and does there come a stage at which it is better to stop and wait for new drugs that are in the pipeline rather than continue with drugs that may be toxic and not very effective? Or is it better to maintain a current regimen – despite virologic failure – if immunological response is maintained, while waiting for new drugs? What to switch to will depend on cross-resistance and resistance testing, knowledge from existing studies, drug interactions, and patient history. Equally important, but less discussed are the practicalities of intensive regimens, including issues such as overlapping toxicities, scheduling, and pill burden.

The use of resistance testing raising a series of important and vexing questions. What to do when there is a mismatch between the patient history and the results of resistance testing? How do we interpret resistance testing in patients who have a long history of anti-viral use, but have been off drug for some time? Should they be re-challenged with drugs they have already taken in order to get a more accurate measure of resistance? Should we be using resistance screening in our inclusion criteria when designing studies?

In tuberculosis, one way to assess new drugs was to take patients with few treatment options, looking at their drug history and the results of their resistance assays. They were matched in pairs and were all given the best therapy they could have and randomized to

receive or not receive a new drug or drugs in addition. This may be a useful strategy in HIV.

We need to look at two types of studies: short term trials to assess whether drugs or combinations have got activity, where we might well reject combinations which don't look effective, and also, long-term studies to assess what the actual impact of the combinations are on morbidity, viral load, CD-4 cell count, etc. Determining appropriate endpoints is another important area for discussion. In trials of heavily pre-treated patients with advanced disease, a sustained response below the limit of detection, though desired, may not be attainable. A stable response above the limit of detection may be a good and more realistic goal. It would be useful to go back to earlier studies to determine if a reasonable viral suppression correlates with differences in morbidity and mortality. Sustained CD+4 cell response may be as important in the longer term if complete viral suppression is not achievable.

Given the large number of possible combination regimens that need to be assessed, factorial and partial factorial trial designs may be particularly useful, because they allow studies to address more than one question at a time.

In addition, if several studies either between groups or within groups (i.e. ACTG and/or CPCRA studies) use a common control arm, there is a useful point of reference for comparison between trials of a broader range of regimens.

We need to discuss whether studies should look at comparison of specific drugs and regimens or comparison of classes of drugs. We also need to sort out issues regarding drug interactions, pharmacokinetics, and toxicities.

Overall, there are different sets of challenges whether one is focused on designing studies to look at long-term risks and benefits or studies to test the activity of new drugs and new combinations. Both kinds of studies are important, but ultimately, it is the longer-term benefit that is important rather than a four-week reduction in viral load.

Jim Neaton, Ph.D.

Dr. Neaton's talk addressed two major issues: the use of resistance testing to define patient inclusion in clinical studies and the issue of generic versus specific drug treatments.

A number of issues need to be considered with respect to resistance testing:

- Cost;
- Logistical factors – real time resistance testing is slow and complex, particularly when doing large multi-site studies;
- Exclusion of people on the basis of specific mutations or phenotypic results limits the ultimate generalized ability of the findings of the study.
- Very specific mutation criteria will require more screening, making studies more costly and slowing up recruitment;
- Misclassification – how many mutations are missing as a result of laboratory error and less specific assays?
- What is the prognostic importance of different mutations?

Data from the GART trial, which was recently conducted by the CPCRA, offer some ideas about how pre-treated patients might be classified for assessment of salvage regimens. In that study, which assessed resistance assays with expert interpretation, patients were divided into several groups:

- No baseline RTI resistance or L90 protease resistance mutation
- No baseline RTI resistance, but L90 protease resistance mutation
- Baseline RTI resistance, but no L90 protease resistance mutation
- Both RTI resistance and L90 protease resistance mutation.

In this study, patients who had one of the markers of resistance did worse than those who had none, while patients who had both markers did worse than those who only had one.

If the results of resistance screening are used as entry criteria for studies, many patients may not be eligible. It might be useful, therefore to have multiple studies – one for those who meet the specific criteria based on a geno- or phenotypic screen, and others for those ineligible for that study.

Salvage studies of specific drugs or regimens are best for short-term study with virologic endpoints because we know that there will be a large group of failures and they may not be kept on the regimen after virologic failure occurs. The best approach may be to test new drugs and/or combination in short-term investigative screening trials and eliminate those regimens that show little or no activity.

To better understand the best long-term strategies, studies involving the comparison of classes of drugs may be more feasible and realistic, though not without difficulty. The heterogeneity of the possible generic treatments is an issue. Are all protease inhibitors alike? If you are doing a trial with specific treatments, it limits your generalizability, and it also makes it much harder to recruit participants for entry into the trial. The wide variety of patients' treatment histories makes it extremely difficult to enroll patients into studies of specific regimens, though they may be willing to enter more generically-defined trials.

Victor DeGruttola, Ph.D.

Screening trials using very short-term endpoints – even one or two weeks – may be useful to determine those candidates that might be appropriate for longer-term trials. But for studies of clinical management, virologic endpoints may not be enough. Long-term follow up is necessary, but only possible when you do not compare specific regimens.

Another major concern is the variability in patients, both in their prior treatment experience and in their genotype. It is important to devise some type of classification system to reduce this variability by using blocking factors in order to compare patients who are more similar. The classification scheme may need to be iterative. Whether using genotype, phenotype or patient history, we would hope that patient will respond very similarly who are within those classes in the studies, but the studies may show us that the classification system needs modification. One way suggested to classify patients is to try and match patients according to some feature of resistance, but that feature still needs to be determined.

Screening studies to test drug activity could be nested into longer-term management trials. If, for example, there was a study comparing a strategy of targeted treatments to a strategy of mega-HAART or a drug holiday; within the targeted treatment arm you could do randomizations of specific combinations, some of which include a new drug.

Janet Anderson, MD

Multi-staged accrual designs in drug screening are studies looking for either any signs of activity for a new drug or a stated level of activity. The designs have been used in cancer clinical trials. They require the ability to classify subjects into fairly homogenous groups and to be able to hypothesize bad success rates and good success rates for a specific group. If you have a very toxic and/or expensive drug, you may hypothesize that you need a greater level of efficacy to warrant further investigation of the drug. The basic design is to accrue a small group and see if the desired level of activity is achieved. If it is, then further accrual can take place. If not, the study is discontinued. The designs are adaptive with a goal of reducing the number of subjects exposed to an ineffective regimen.

For instance, Dr. Anderson described a trial in which patients would be treated with a nucleoside backbone (2 NRTIs), and would then be briefly randomized to add either an NNRTI or a protease inhibitor. If both were active following a brief period, then patients could be allowed to combine the NNRTI and the protease inhibitor with the nucleoside backbone. Intensive pharmacokinetics could be performed to ensure that the combination resulted in appropriate drug exposure levels, and all patients could be followed for long-term outcome. In addition, Dr. Anderson proposed, patients might be randomized between different nucleoside backbones.

CHOICE OF ENDPOINTS IN SALVAGE THERAPY TRIALS

The goal of this panel was to discuss issues regarding the appropriate measures of activity and efficacy for antiretroviral therapies in patients for whom earlier regimens have failed.

Dan Kuritzkes, MD

In a presentation regarding choice of endpoints in trials of heavily pre-treated patients, Dr. Dan Kuritzkes pointed out that the usual endpoint in recent HIV therapeutic trials has been the percentage of patients who have achieved undetectable plasma HIV RNA levels. However, a number of recent studies have suggested that, in pre-treated patients, success rates are very low. Dr. Kuritzkes showed data supporting the use of change in plasma HIV RNA over significant periods of time as a surrogate marker for clinical efficacy. In a number of trials, at least a half-log reduction in plasma viral load over a year or two has translated into significant clinical improvement. Therefore, Dr. Kuritzkes proposed, that a strict focus on the percentage of patients who achieve undetectable RNA levels may obscure other important benefits of anti-HIV therapy.

Potential candidates for study endpoints include:

- Time to treatment failure (although the definition of treatment failure remains problematic)
- Change in plasma HIV RNA from baseline
- Composite endpoints which measure changes in the CD4 Cell Count, RNA level, and occurrence of clinical events

Dr. Kuritzkes proposed that, ideally, trials would measure both short-term and longer-term changes in plasma HIV RNA levels as the primary endpoint, and would then measure CD4+ cell counts, percentages of patients with undetectable RNA levels, and clinical events as secondary endpoints.

Importantly, Dr. Kuritzkes also emphasized the importance of distinguishing studies that test the efficacy of individual therapies from studies that measure the utility of a treatment regimen. While the former are necessary for drug approval, the latter are more important for the clinical management of heavily pre-treated patients.

Douglas Manion, MD

In a presentation by Dr. Doug Manion, he pointed out that the choice of endpoints should be coupled with the question of which analyses are appropriate. In his company's study, DMP-06, which compared AZT/3TC/indinavir to AZT/3TC/efavirenz and indinavir/efavirenz, patients were treated for forty-eight weeks. The study was analyzed using two different methods: one method censored patients after discontinuation of study drug, and the other analysis counted discontinuations as treatment failures. The findings were different, though, and as Dr. Manion pointed out, "almost certainly the biological truth in terms of relative activity of compounds or regimens is somewhere between these bookends."

Dr. Manion also pointed out that not all study discontinuations are the same. Although, in DMP-06 thirty-five percent of patients were not on-study and/or fully suppressed by week forty-eight, only fifteen percent had experienced virologic failure or severe toxicity, and the remaining twenty percent had dropped out due to other reasons.

Richard Pollard, MD

Dr. Pollard presented data on markers of immunologic benefit from therapy that could be used in clinical trials. He first summarized the natural history of the immune system, and the effects of HAART on the system. In HIV-infected patients, total CD4+ cell counts decline over time, including both memory and naïve cells. Levels of activated CD38+ T-cell rise. Following treatment, there is a relatively rapid rise in total CD4+ cells, with a rapid increase in memory cells, and a somewhat slower and more prolonged rise in naïve cells, as well as a decrease in activated cells. In addition, pathogen-specific immunity may begin to regenerate following successful HAART therapy. While total CD4+ cell counts have risen continuously throughout 2 - 3 of treatment in a number of studies, and such measurements are clearly useful, he said, data are accumulating which may allow the use of other markers of immunologic success or failure in clinical trials.

However, Dr. Pollard warned, there are therapies – hydroxyurea, for example, that improve the virologic response while blunting the CD4+ response. "I think one of the

major issues that we all have to come to consensus about," he said, "is what to do with patients that are virologic successes, but don't have changes in CD4+ count." Should we change anti-retroviral regimens when there is no CD4 response? Dr. Pollard suggested that this might depend on the baseline CD4 count – patients with counts <200 might consider changing therapy, while patients with higher counts would not. We need to define the threshold of what might be considered a successful CD4 response. Is it a 50 cell increase? Or a 100 cell increase? Should the goal be to go over 200 CD4? Conversely, in patients who are succeeding virologically, but have falling CD4 counts, what is considered failure – a drop in 50, 100 or more cells? We do not have answers to these questions.

Stephen Deeks, MD

Dr. Deeks presented data demonstrating sustained CD4+ responses and clinical improvement in patients treated with protease inhibitors, despite virologic failure. Data from a Swiss cohort showed no difference in the risk of developing AIDS or death between people who maintain undetectable viral load and people who rebounded after reaching an undetectable level. Such a response could be due to the decreased fitness of resistant virus. "In the setting of nucleoside antiviral therapy and probably even the successful virologic response to a protease-inhibitor base regimen, HIV RNA levels…serve as a good surrogate marker," he said. "The question I think we still need to struggle with is to what degree HIV RNA levels predict a poor response in the setting of virologic failure."

With regards to whether clinical endpoints should be used in studies of heavily pre-treated patients, Dr. Deeks felt it would be impractical. The studies would need to be very large, or initially directed at patients with very advanced disease. Also, many patients in his cohort have died of causes not traditionally thought to be HIV related, e.g. cancers or cardiac-related. This raises an important question: does anti-retroviral therapy, even if no longer successful, change the outcome and clinical manifestations of HIV disease?

Jeff Murray, MD

Dr. Jeff Murray of the US Food and Drug Administration (FDA) presented the agency's perspective on the choice of endpoints in efficacy trials. As general guidance, according to Dr. Murray, the FDA recommends that companies seeking accelerated approval for anti-HIV therapies show data on the percentage of patients below the limit of detection at twenty-four weeks. However, he also noted that "we're not that rigid. Mean changes from baseline might be acceptable in certain circumstances."

Dr. Murray cited the example of a study comparing two nucleosides administered with either the Invirase hard-gel capsule formulation of saquinavir or the Fortovase soft-gel capsule formulation over sixteen weeks. The regimens were comparable in terms of mean change in plasma HIV RNA levels at 16 weeks. However, patients taking Fortovase were more likely to sustain undetectable plasma RNA levels after 16 weeks. In addition, Dr. Murray noted that use of time to failure in patients with undetectable plasma HIV RNA levels "protects the treatment comparison," because the endpoint occurs at the same time that treatment would usually be changed.

Dr. Murray gave several examples of efforts to correlate virologic response with clinical outcome, and suggested that a 0.3 log reduction in plasma HIV RNA levels always resulted in clinical improvement if sustained beyond eight weeks. Smaller reductions, or reductions that were lost prior to the eight-week period sometimes resulted in clinical benefit, but not uniformly.

In response to suggestions that lower hurdles be used, such as proportion of patients with less than 1,000 or 5,000 copies of virus/ml, Dr. Murray labeled such choices "arbitrary," and suggested that "It would probably be best to stick with something that means something biologically."

In summary, Dr. Murray proposed that, when use of the "proportion of patients with undetectable viral load" endpoint was not feasible, changes in plasma viral load could be

used. Such changes, however, needed to be sustained over a longer period of time – such as twenty-four weeks --in order to suggest clinical significance. Still, he warned that, due to selective dropout, changes in viral load are difficult to interpret at even longer time points, such as forty-eight weeks.

Discussion

During the discussion, participants commented on the difference between registrational trials, which attempt to measure the utility of particular agents, and longer-term strategy trials, which measure the effectiveness of more generic regimens or treatment approaches. All participants recognized tensions between these two approaches, however there was some disagreement as to whether these tensions could be resolved by nesting small short-term studies into larger strategy trials.

One study proposal focused on strategic therapeutic interruptions, in which HAART regimens, providing either full or partial suppression, would be alternated with structured drug holidays with intensive monitoring of immunologic and virologic parameters. Another approach proposed was the use of regimens such as ddI/d4T/HU, which would provide some viral suppression, without significant risk to future regimens.

One participant proposed a model used in the evaluation of cancer therapies, which he dubbed "winner takes all." A short-term screening trial is conducted, comparing the activity of a variety of different therapies and/or regimens. After a few weeks, the most potent treatments are selected, and patients are then randomized between those options for longer-term evaluation.

Another participant made the point that clinicians see heterogeneity among treatment failures. Some patients are failing dramatically despite potent therapy, and may require drastic intervention, while some patients have minimal ongoing viral replication. Therefore, inclusion criteria for salvage therapy trials need to be carefully considered, depending on the interventions being offered.

The question of appropriate endpoints was argued extensively. One participant noted that "there are multiple toxicities involved, and…we've got to get beyond AIDS-defining progression of disease and look at all-cause mortality." Another participant added the idea that, with increases being noted in diseases that were not classic AIDS-defining illnesses, the natural history of the disease may have changed in ways that wouldn't be captured by standard clinical endpoint definitions. The importance of long-term follow up, even in patients who have discontinued study drug, was extensively supported by participants.

PHARMACOLOGY PANEL

This panel addressed the pharmacology portion of the workshop, which considered "Pharmacological Issues Relevant to Treatment Failure, Adherence, Drug Interactions, Drug Uptake, Distribution, and Therapeutic Drug Monitoring".

Terry Blaschke, MD

Dr. Blaschke focused primarily on adherence, drug exposure and their connection to drug failure. He noted that long-term adherence to HAART and mega-HAART is not feasible for many people with HIV and he posited the question, "How much adherence is enough?" Many people who fail therapy do not have phenotypic or genotypic resistance to their drugs, which might imply that drug exposure is insufficient to express resistance or suppress viral replication. Drug exposure is a function of two specific variables: a person's pharmacokinetics, and a person's drug-taking behavior. Therefore, drug exposure is the product of these two variables. When we talk about what is going to happen to the virus, we have to think of it in terms of drug exposure and not just in terms of adherence, pharmacokinetics or drug interactions.

There is a clear relationship between drug exposure, maximum response, adherence and the probability of developing a drug resistant virus. It is important to understand drug exposure, adherence and pharmacokinetics because they can help explain why someone

failed a particular regimen and use that information to put together a new regimen for a greater chance of success.

There is a need for observational studies and a considerable amount of effort needs to be put into collecting all of the potentially significant variables in these studies, including adherence, in order to generate information to be used in the larger prospective studies.

Dr. Blaschke concluded his presentation with a number of questions. What do we know and what do we need to know about adherence and treatment failure? How should we measure adherence in clinical trials? What do we really need to know about adherence? Do we need to know the average amount of drug taken? Do we need to know the variability in dosing times? Do we need to know the occurrence and length of drug holidays? Do we need to know the duration of time during which the drug concentrations were above or below the IC90 or IC50? Dr. Blaschke suggested that we need to know all of the above.

Alastair Wood, MD

Dr. Alastair Wood from Vanderbilt University described the multi-drug resistance gene, which is the gene that codes for the drug transporter known as p-glycoprotein. Drugs normally enter cells either by a passive process of diffusion or they may be actively taken up into cells (which is the case for many protease inhibitors), and then they are out of the cell. Some of the drugs are pumped out by p-glycoprotein. Both nelfinavir and saquinavir are transported out of cells that express p-glycoprotein.

P-glycoprotein is expressed in a number of tissues. It is expressed in the intestines, liver, kidney, testes and lymphocytes. As p-glycoprotein pumps drugs out of the cells, drug absorption may be reduced. P-glycoprotein is also expressed in capillary and epithelial cells that make up the blood-brain-barrier. Mice experiments have shown that nelfinavir concentrations in the brain are significantly higher (about 40-fold higher) in p-glycoprotein knock-out mice than in wild type mice. Similar results have been seen with

indinavir and saquinavir. Studies have also shown that drug levels are higher in the testes and other tissues in p-glycoprotein knock-out mice compared to wild type mice.

One of the difficulties is that there is considerable overlap between substrates for p-glycoprotein and cytochrome p450 3A4, one of the enzymes which is responsible for metabolism for a number of protease inhibitors, so a lot of these drugs are inhibitors of both. A drug can be pumped out a cell before it is metabolized by an enzyme like cytochrome p450 3A4 or in the case of a metabolite, after it has been metabolized by 3A4 and because of the considerable overlap between p-glycoprotein and the cytochrome p450 3A4 substrate inhibitors. These need to be considered when evaluating the efficacy of these therapies.

During drug discovery, we might want to look for drugs that are potent inhibitors of p-glycoprotein, to increase drug entry into sanctuary sites, while having less potent inhibition of cytochrome p450 3A4.

If we are able to develop a potent p-glycoprotein inhibitor, we may be able to increase drug entry into cells. This is important for a number of reasons. Not only would you be able to get higher concentrations of drug into the brain, but you may also be able to get higher drug concentrations into cells at lower doses of drug, which might also reduce drug side effects and greater efficacy.

John Gerber, MD

Dr. John Gerber from the University of Colorado discussed the important of drug-drug interactions. When talking about drug-drug interactions, we need to define whether the interaction is pharmacodynamic or pharmacokinetic. It is easier to just talk about pharmacokinetics since we can measure drug levels. It is a simple way of looking at what the body does to a drug as it is being administered – absorption, metabolism, distribution, protein binding etc. Pharmacodynamics on the other hand is how the body (in this case the virus) handles or interacts with the drug.

It is very difficult to evaluate pharmacodynamics *in vitro* because of the complexity of what happens *in vivo*. As a result *in vitro* experiments that show synergy do not necessarily have *in vivo* correlation.

The synergy between protease inhibitors and reverse transcriptase inhibitors may be easier to explain. The protease generates more reverse transcriptase and *vice versa*. Reverse transcriptase is needed to generate more virus so there could be a potential interaction.

It is more difficult to explain the synergy between the protease inhibitors. For instance Abbott has done studies showing synergy between ritonavir and saquinavir and between ritonavir and indinavir, however Dr. Martin Hirsch has shown some antagonism between indinavir and saquinavir *in vitro*. We do not understand the mechanism for these interactions and it is important to think about this in the future especially since ritonavir inhibits cytochrome p450 3A4 but also appears to inhibit p-glycoprotein also. Synergy may also be occurring inside the cell.

There are also genetic mutations which result in synergy such as the 3TC-related 184 mutation. There may be many more mutations that we do not really understand from a genetic standpoint, so that when drugs are used by themselves they may be potentially resistant, but still somehow result in some response when used as part of a combination regimen.

We are much more comfortable in enhancing pharmacokinetic interactions since we can measure drug levels. Efficacy can be increased because you can raise either the trough levels or the area under the curve. Pharmacokinetics can also be improved so that drugs do not have to be dosed as frequently. We generally have a good idea about two-drug interactions, we know much less about three-drug interactions and virtually nothing about four- and five- drug interactions.

In ACTG 359, for example, an unpredicted interaction was found between adefovir dipovixil and delavirdine. Based on drug metabolism, there would have been no way to predict that adefovir would affect the kinetics of delavirdine, which subsequently affected the kinetics of saquinavir and ritonavir. There are going to be very complex interactions when multiple drugs are used not just from pharmacokinetic interactions but sometimes from unwanted pharmacodynamic interactions. For instance, with ritonavir and indinavir, a popular combination, indinavir is a substrate for cytochrome p450 3A4 and ritonavir is a potent inhibitors of p450 3A4. This results in a decrease in indinavir dose and a ten to twenty fold increase in trough levels, compared to the standard dose of indinavir. But, we do not know what to expect when three-drug combinations, such as ritonavir, indinavir and delavirdine are used, since delavirdine increases both indinavir and ritonavir levels. Similarly we do not know what to expect with ritonavir, saquinavir and efavirenz since efavirenz significantly decreases saquinavir levels but ritonavir significantly increases saquinavir levels. Besides understanding the importance of drug/drug interactions we also need to understand the therapeutic index.

We need to understand which drug concentrations should be increased. Ritonavir, for instance, is not a drug that you want to increase the concentration of, since very few people will be able to tolerate increased concentrations.

There are also some drugs that, no matter how much the dose is increased, you will not see greater activity. One example is the nucleoside analogues, because probably not much more triphosphate metabolites can be produced.

Craig Hendrix, MD

Craig Hendrix from Johns Hopkins discussed the importance of conducting intensive drug interaction studies for both pharmacokinetics and pharmacodynamics.

It is important that interaction studies be conducted when drugs are at steady state. Pharmacokinetic studies that provide an accurate assessment of the interaction often cannot be done with single doses. It can take ten to fourteen days for some of the drugs to

reach steady state. This adds quite a bit of complexity if the interaction study is the lead-in to a longer-term clinical study. It certainly adds a degree of risk when a single drug is added to a combination regimen. This may be acceptable in a population that is virologically well-controlled, but it is much more risky in the salvage population.

Most pharmacologists would like to do studies where you have the pharmacokinetics of drug A, then drug A and B and then drug B alone and compare the different drug levels.

It is difficult to conduct these studies because clinicians and patients alike are concerned about the possibility of developing drug resistant virus. For instance, people on ritonavir and saquinavir who are well suppressed (less than 50 copies HIV RNA per milliliter) are afraid to add efavirenz because of the possibility of developing resistance to efavirenz even though they are only on drug for two weeks.

It is likely that these studies can be done in healthy volunteers, as they are very good predictors of the kinetics that you will see in an HIV-infected individual. However, there is a theoretical possibility that the immunological stimulation as a result of HIV replication may have an effect on cytochrome p450 enzymes and other systems. These will generally be inhibitory effects so drug levels should increase and may actually be helpful. However, in some small studies and animal studies there have also been induction effects.

A traditional intensive pharmacokinetic study involves about a dozen patients at each dose level whose drug levels are measured at about a dozen different time points. Since drug levels are measured at the same time among all of the patients, you get a good idea of the half-life of the drug (based on the clearance and the volume distribution) and the variability of the area under the curve.

A population pharmacokinetic study usually involves a lot of patients with many fewer measurements per patient and can still answer many questions with a fair degree of precision. This type of study is more convenient for the patient and logistically may be

easier to do as sub-studies. A number of co-variables, such as age, gender, disease stage, concomitant medications, renal and hepatic function, can also be looked at. However, with each variable assessed, one would need more patients.

Not every site can do intensive pharmacokinetic studies, as an in-patient unit and staff that are specially trained in this area are required. Population pharmacokinetic studies, on the other hand, are relatively simple. They require blood draws that are sent to a central laboratory. The data analysis, however, is much more complex in the population pharmacokinetic studies.

Ideally a correlation can be made between drug levels and changes in viral load or immune function. Then it may be possible to aim for that drug level when treating patients. However, it is necessary to understand the relevant variables in advance. Are the peak levels or trough levels more important? Or is it the area under the curve? Or is it the relationship between the peak and trough ratio to the IC90?

It may be possible to study some of this in *in vitro* experiments. Based on the hollow fiber model, it may be possible to measure how much virus is still being produced after introduction of a drug and it may be possible to determine how quickly resistance develops to a particular drug. This information may be used to optimize treatment regimens to be used in longer term treatment studies.

It is probably useful to set up a data collection system within the clinical trial networks to gather data on demographics, creatinine clearance, hepatic function, and blood samples to measure drug levels. These can be sent to a central laboratory and analyzed providing more informative data on drug interactions among the heavily pretreated population.

Gene Morse, PharmD

Dr. Gene Morse from SUNY Buffalo discussed the pros and cons of therapeutic drug monitoring.

Most people familiar with therapeutic drug monitoring believe that if drug concentrations are measured and the dose of the drug is adjusted upwards or downwards, there will be better activity or reduced toxicity.

One area of concern is around clinical needs versus regulatory needs versus pharmaceutical company needs. Most of the drugs that are available on the market are based on a fixed dose regimen after going through dose escalation studies and multiple dose studies etc. However, this may not be the best approach today.

It might be reasonable to try and achieve a certain drug concentration since the non-nucleosides and the protease inhibitors have tremendous inter-patient variability. But it is also important to know the intra-patient pharmacokinetics so that if there is a dose adjustment, there would be minimal variation in drug concentration over time.

HIV susceptibility testing may also help in the treatment strategy approach. For instance if someone is virologically failing a regimen, such testing can help determine if someone was actually taking the drug. If so, what kind of drug exposure was achieved? With a phenotypic test, it is possible to identify the IC90 of the predominant virus species and determine if a higher dose is needed. It is entirely likely that the doses required to get above the IC90 to achieve viral suppression are higher than the approved doses of a drug or even higher than dose studies in the drug development process.

It might be possible to use therapeutic drug monitoring to vary the intensity of drug exposure in relation to the IC90 value. It may be possible to alter the dose of various drugs within a combination rather than switching to a whole new regimen.

It is possible to measure drug levels in real time in clinical studies. This is, of course, very labor intensive. For example, if one of the drugs in a combination is a known inducer, it is possible to measure its effects on the other drugs in real time, and takes three to five days.

Another area for therapeutic drug monitoring is where people have different absorption and metabolic rates. Not everyone is at the same stage of disease progression. People who are on their fourth or fifth regimen are likely to have more advanced stage disease progression. There are a number of people who have achlorhydria, while others have concurrent hepatitis. Therapeutic drug monitoring may be helpful in optimizing antiviral effects and/or minimizing the development of toxicities.

One very important issue in therapeutic drug monitoring is to ensure that the laboratories have a QA/QC program so that all of the data can be used.

Discussion

A recommendation was made to establish a consensus group to focus on standardization of methodologies for blood collections, validation of assays, and issues around drug interaction when steady state levels are achieved. There should also be QA/QC standards across the labs.

One of the big problems with strategy trials is that there are no pharmacokinetic or pharmacodynamics data for most of the commonly-used salvage regimens. It is important to start collecting some population pharmacokinetic data. Ideally, population pharmacokinetic sampling can be built into expanded access programs rather than doing pharmacokinetic/pharmacodynamic studies in people on a defined regimen.

In the clinics, "experiments" are taking place everyday since some people are also on anti-seizure medications, methadone and psychiatric medications and it may not be known whether these drugs effect the antiretrovirals or vice versa. One area of focus is to get the pharmaceutical companies to work together to collect drug interaction data early in the drug development process. Since most researchers/clinicians know what drugs they would probably use with the new drugs, it is critical that this information is available early on. Otherwise, we will continue to put people on these combinations blindly and only later find out that they developed resistance to the new drug or to that class of drug

because of an interaction. It is possible to prioritize some of these interactions based on their metabolic pathways but it is also important to consider unexpected interactions.

Another recommendation is to have a clearinghouse for drug interactions. Additionally the pharmacokinetic substudies should not be analyzed at the end of the large study because by then it is too late and one group may have already unnecessarily failed their regimen.

Another recommendation is to make sure that the assays to measure drug levels are available to people who want to do pharmacokinetic studies. Presently a lot of these assays are not available to the public because they are considered proprietary. Some companies will not release pure samples of their compound for investigators who wish to set up an assay. A group, consisting of people from industry, academia, government and community, should be formed to address the questions that need to be answered in a particular protocol. They would make sure that the assays are ready and that the turn around time is not too long. While the majority of the blood samples still have to assayed, there will be preliminary information to know what needs to be done in the short and long term.

Another recommendation is for the FDA to make sure the assays used by the company become available once a drug is approved. Often, the companies use an outside laboratory and they will not release any information on the assays. On the other hand, many universities also want royalty rights when an assay is developed.

REGULATORY AND INDUSTRY ISSUES

This panel discussed the needs of industry in developing new therapies and therapeutic regimens for the treatment of heavily pre-treated patients.

Heidi Jolson, MD

Dr. Jolson discussed the regulatory issues that arise in the context of using experimental agents in heavily pre-treated patients. In general, she said, the agency has a lot of

flexibility about designating a condition as "serious or life-threatening." The agency has accepted the principal that risk-benefit judgements are necessary, and that patients and providers are willing to deal with greater uncertainty regarding the benefits of therapy when the risks of untreated illness are high.

Dr. Jolson described the various mechanisms by which FDA can make experimental agents available to patients for use in salvage regimens:

1) Randomized, controlled trials
2) Emergency INDs, in which approval for emergency access may be granted by phone in life-threatening situations with twenty-four hour turnaround.
3) Open-label protocols, such as Treatment INDs and Parallel Track Programs, which should collect safety data, and may collect limited efficacy data.

The mechanism used to provide access may depend on particular circumstances and available data. For instance, single patient access may be granted based on limited *in vitro* activity data, whereas broader access generally requires some clinical safety and efficacy data.

According to Dr. Jolson, FDA does not object to the use of multiple experimental agents in any of these settings, although the rules for data collection may differ somewhat. Single-patient use may require only basic safety data collection, while open-label protocols generally require more systematic safety data collection, and may include some limited assessment of efficacy.

Dr. Jolson also noted that, when multiple investigational agents are used in a registrational trial, investigators need to be able to assess "the relative contribution of each individual drug regarding safety and efficacy."

Jeff Chodakewitz, MD

Dr. Chodakewitz commented on the importance of distinguishing registrational trials from access programs. In his discussion, he chose to focus on the conduct of clinical trials.

From Dr. Chodakewitz's perspective, companies have to consider a number of factors when deciding to initiate salvage therapy studies. Most important, he stated, is a preliminary determination of the product's potential for success in a salvage therapy setting. For instance, a new drug that is ineffective against the M184V mutation is probably not appropriate for salvage in patients who have previously been treated with lamivudine.

In addition, there are a number of considerations for crossing what he called "the risk threshold," including:

1) Pharmacokinetic properties of the product
2) Safety profile
3) Impact of different baseline resistance mutations
4) Potential consequences of therapy for future drug resistance

Of course, companies developing a new drug must consider the impact of salvage therapy trials on what is usually a very limited drug supply.

Dr. Chodakewitz also noted that companies and regulators need to consider the goals of salvage therapy. Given the low likelihood of success with current regimens, the goal of undetectable viral load at twenty-four or forty-eight weeks may be unachievable, and other measures of product efficacy may be more appropriate.

He also pointed out that the potential for success is important. While negative results can be very important from the perspective of the patient and clinician, when such results are obtained early in the course of development, they may shape the general perception of the product in ways that are harmful to the company's future regulatory and marketing goals.

Finally, while Dr. Chodakewitz endorsed the idea of using multiple experimental agents in salvage therapy regimens, he pointed out that the FDA standard of distinguishing the relative contribution of each agent presents vexing trial design issues for companies developing new products.

Alex Dusek, MD

Dr. Dusek discussed the challenges to a small company of conducting trials in heavily pre-treated patients. For a company with limited resources, it is necessary to decide whether to pursue such studies in registrational trials, or in Phase IV post-marketing studies. Dr. Dusek said that companies are often trying to meet the needs of diverse audiences with limited financial resources and drug supply. For instance, he said, investors frequently do not understand the urgent need for salvage therapies due to a misperception that few patients experience treatment failure.

In addition, trial design issues, particularly in collaborative efforts, can be especially troublesome. For instance, Dr. Dusek said, the absence of an accepted standard of care for heavily pre-treated patients presents particular difficulties in designing controlled salvage therapy trials. Other issues that arise in these collaborative efforts include:

1) The need for pharmacokinetic studies
2) The different rates at which manufacturers scale-up product supply
3) The need to achieve consensus on protocol designs
4) The need for rigorous confidentiality agreements.

Dr. Dusek suggested that, at least for smaller companies, the overriding concern is the urgency of obtaining product approval.

In his opinion, the current treatment guidelines structure, which does not distinguish between particular therapies or regimens with respect to utility in salvage regimens, provide no incentives for manufacturers to study such regimens. Developing such incentives should be a priority. Some suggestions offered by Dr. Dusek include specific labeling language regarding salvage utility, and establishment of an efficacy standard that might differ from that applied to treatment for naïve patients.

Finally, Dr. Dusek discussed several trial design models that might reconcile the regulatory requirements with the demands of patient care.

The standard trial design of Standard of Care (SOC) + New Drug X vs. SOC + placebo is probably not acceptable due to the failure of current standard regimens to control virus in heavily pre-treated patients, and the need to treat patients with at least two active drugs. Therefore, Dr. Dusek proposed an alternative trial design of:

- SOC + New Drug X + New Drug Y vs.
- SOC + New Drug X + Placebo

However, this still leaves a treatment arm (Arm B) with only one new drug. Therefore, Dr. Dusek proposed that optimal design might be:

- SOC + New Drug X + New Drug Y vs.
- SOC + New Drug Z + New Drug Y vs.
- SOC + New Drug X + New Drug Z

While probably the most clinically optimal (or potentially generous) of the three trial designs, the acceptability of the last study for regulatory purposes is unclear.

Roy Gulick, MD

Dr. Gulick discussed the clinical care issues that arise in designing salvage therapy studies. He shared a story about a patient who, with the best available clinical care, had rapidly cycled through all available anti-HIV therapies, including several experimental drugs, and had no new treatment options. According to Dr. Gulick, there is little potential for success of current drugs in salvage regimens. There are a variety of conflicting paradigms that researchers are trying to sort out:

1) Regulatory demands versus the demands of clinical care
2) The need to treat individuals versus the need to assess therapeutic effects
3) The need for clinical strategy trials versus the need to assess specific therapies

The patient population is heterogeneous with respect to previous treatment history, current immunologic, virologic and clinical status, and presence or absence of baseline resistance mutations. There are multiple possible regimens to be studies, and, Dr. Gulick reiterated, no generally accepted control arm.

In combining experimental agents for salvage therapy studies, there are multiple difficulties:

1) Distinguishing the cause of adverse events
2) Defining drug interactions
3) Defining pharmacokinetics
4) Determining resistance patterns

In addition, Dr. Gulick raised the question of the appropriate forum for conducting such trials. At present, there are a number of possibilities, including the Inter-company Collaboration (ICC), the AIDS Clinical Trials Group (ACTG), the Community Program for Clinical Research on AIDS (CPCRA), other NIH programs, independent pharmaceutical collaborations, and other possible venues.

Discussion

During the discussion, workshop participants agreed that market forces are pushing companies to conduct salvage therapy studies. As one participant suggested, there is a competitive advantage to collaboration.

Participants picked up on the suggestion raised by Dr. Dusek of the need for regulatory innovation to offer companies incentives to participate in salvage therapy studies. However, another participant cautioned that increased regulatory requirements would likely result in delayed marketing time.

One workshop attendee pointed out that studies of new products in salvage regimens may ultimately be more generalizable than studies in naïve patients; one can probably extrapolate success in a salvage setting to the naïve setting, while the reverse may not be true.

Finally, several conference participants worked together to produce a study proposal that could evaluate several experimental agents at the same time, while offering study participants a reasonable chance of therapeutic success.

1) SOC + Drug X + Drug Y + Drug Z
2) SOC + Drug X + Drug Y
3) SOC + Drug X + Drug Z
4) SOC + Drug Y + Drug Z
5) Drug X + Drug Y + Drug Z

In addition, the study designers noted that such a study would make very efficient use of a limited number of patients.

The Challenges of Clinical Trial Design in Assessing the Effects of Anti-HIV Therapy in Heavily Pre-treated Patients

The goal of this meeting is to discuss issues in the design and implementation of studies of salvage therapy regimens in heavily pre-treated patients. The meeting will allow for the presentation of differing needs, priorities and challenges faced by industry, researchers, regulators and patients conducting and participating in this research. At the meeting, we will try to agree upon some definitions for treatment failure and success. Further, we will work to develop a better understanding and agreement between parties regarding what is necessary and feasible when designing studies of new drugs for indication in salvage therapy.

WORKSHOP AGENDA

Friday, May 21

4:00	**Introductory Remarks**	
	Roy Gulick, M.D.	Cornell University
4:10	**Welcome**	
	David Barr	Forum for Collaborative HIV Research
	Ben Cheng	Project Inform
	Spencer Cox	Treatment Action Group
	Bill Duncan, Ph.D.	NIAID Division of AIDS

Challenges in study design: five perspectives

4:15	**Industry perspective**	
	Franck Rousseau, Ph.D.	Triangle Pharmaceuticals
4:30	**Regulatory perspective**	
	Heidi Jolson, Ph.D.	FDA Division of Antiviral Drug Products
4:45	**Virology perspective**	
	John Mellors, M.D.	University of Pittsburgh Medical School
5:00	**Clinical perspective**	
	Julio Montaner, M.D.	Canadian HIV Trials Network
5:15	**Patient perspective**	
	Carlton Hogan	CPCRA Statistical Center
5:30	**Overview / Discussion**	
	Roy Gulick, M.D.	Cornell University

Saturday, May 22

Statistical issues in trial design – Issues will include selection of controls when studying multiple drugs in a population with advance disease, including a discussion of other disease models.

Panel co-facilitators:

Spencer Cox	Treatment Action Group
Victor DeGruttola, Ph.D.	Harvard University School of Public Health
Janet Andersen, M.D.	Harvard University School of Public Health

Panel:

Janet Darbyshire, M.D.	University College London Medical School
Daniel Kuritzkes, M.D.	Univ. of Colorado Health Sciences Center
Jim Neaton, Ph.D.	University of Minnesota

8:00 Panel presentation

8:30 Discussion

9:45 Coffee Break

Defining treatment "success" and "failure" - Issues will include attempting to define treatment "success" and "failure" in a heavily pre-treated population, while discussing the determinants used to evaluate when a "failing" regimen should be switched.

Panel co-facilitators:

David Barr	Forum for Collaborative HIV Research
Carla Pettinelli, M.D.	NIAID Division of AIDS

Panel:

Daniel Kuritzkes, M.D.	Univ. of Colorado Health Sciences Center
Richard Pollard, M.D.	University of Texas Medical Branch
Doug Manion, M.D.	DuPont Pharmaceuticals, Inc.
Jeff Murray, M.D.	FDA Division of Anti-viral Drug Products
Stephen Deeks, M.D.	University of California at San Francisco

10:15 Panel Discussion

10:45 Discussion

12:00 Lunch

Pharmacological issues relevant to treatment failure: Adherence, drug interactions, and therapeutic drug monitoring -Issues will include drug-drug interactions, drug uptake, challenges of delivering drug to sanctuary sites and adherence to mega-HAART regimens.

Panel co-facilitators:

Ben Cheng	Project Inform
Terry Blaschke, M.D.	Stanford University School of Medicine

Panel:

Alastair J.J. Wood, M.D.	Vanderbilt University
Craig Hendrix, M.D.	Johns Hopkins Univ. School of Medicine
John Gerber, M.D.	Univ. of Colorado Health Sciences Center
Gene Morse, Pharm.D.	State University of New York at Buffalo

1:00 **Panel Presentation**

1:30 **Discussion**

2:45 **Coffee break**

Perspectives of access to multiple therapies in combination salvage protocols - Issues will include access to multiple drugs for combination salvage therapy studies.

Panel co-facilitators:

Bill Duncan, Ph.D.	NIAID Division of AIDS
Roy Gulick, M.D.	Cornell University

Panel:

Heidi Jolson, Ph.D.	FDA Division of Antiviral Drug Products
Alex Dusek, M.D.	Trimeris Pharmaceuticals
Jeff Chodakewitz, M.D.	Merck Research Laboratories

3:15 **Panel Presentation**

3:45 **Discussion**

5:00 **Summary of Workshop**
Roy Gulick, M.D. Cornell University

The Challenges of Clinical Trial Design in Evaluating HIV Antiretroviral Use in Heavily Pre-treated Patients

Participant List

Janet Andersen, M.D.	Harvard University School of Public Health
Raffi Babakhanian	European Community Advisory Board
David Barr	Forum for Collaborative HIV Research
John Baxter M.D.	Cooper Hospital University Medical Center
Mark Becker, Pharm.D.	Agouron Pharmaceuticals, Inc.
Terrence Blaschke, M.D.	Stanford University
Carol Brosgart , M.D.	Gilead Sciences, Inc.
Scott Brun, M.D.	Abbott Laboratories, Inc.
Richard Carroll, Ph.D.	Visible Genetics, Inc.
Ben Cheng	Project Inform
Henry Chang, M.D.	DuPont Pharmaceuticals, Inc.
Jeffrey Chodakewitz, M.D.	Merck Research Laboratories, Inc.
Spencer Cox	Treatment Action Group
Janet Darbyshire, O.B.E., F.R.C.P.	University College London Medical School
Victor DeGruttola, Ph.D.	Harvard University School of Public Health
Nikos Dedes	European Community Advisory Board
Lynda Dee	AIDS Action Baltimore
Stephen Deeks , M.D.	San Francisco General Hospital/UCSF
Bopper Deyton, M.D.	U.S. Department of Veterans Affairs
William Duncan, Ph.D.	NIAID Division of AIDS
Alex Dusek, M.D.	Trimeris, Inc.
Wafaa El-Sadr , M.D.	Harlem Hospital
Laurent Fischer, M.D.	DuPont Pharmaceuticals
Caroline Fortier, Ph.D.	BioChem Pharma
Brian Gazzard, M.D.	Chelsea and Westminister Hospital
John Gerber, M.D.	University of Colorado Health Sciences Center
Diane Goodwin, Pharm.D.	Glaxo Wellcome, Inc.
Neil Graham, M.D.	Glaxo Wellcome, Inc.
Linda Grinberg	FAIR
Roy Gulick, M.D.	Cornell University
Mark Harrington	Treatment Action Group
Fred Harris, M.D.	Boehringer Ingleheim
Peter Hawley, M.D.	Agouron Pharmaceuticals, Inc.
Nicholas Hellman, M.D.	Virologic, Inc.
Jerry Helton	Pharmacia & Upjohn
Craig Hendrix, M.D.	Johns Hopkins University
Geoffrey Henson	AnorMED Inc.
Carlton Hogan	University of Minnesota
Heidi Jolson, M.D.	FDA
Bertil Jolson, M.D.	European Medical Products Agency
Dan Kuritzkes, M.D.	University of Colorado Health Sciences Center
Kiyoshi Kuromiya	Critical Path AIDS Project

Joep Lange,M.D.	Amsterdam Medical Centre
Jody Lawrence, M.D.	ARAC
Francois Lebel , M.D.	Parke-Davis Pharmaceutical Research, Inc.
Douglas Manion, M.D.	DuPont Pharmaceuticals, Inc.
Douglas Mayers, M.D.	Henry Ford Hospital
Colin McLaren, M.D.	Bristol-Myers Squibb, Inc.
John Mellors, M.D.	University of Pittsburgh Medical Center
Michael Miller, M.D.	Gilead Sciences, Inc.
Veronica Miller, M.D.	Klinikum der J.W. Goethe University
Julio Montaner, M.D.	University of British Columbia
Gene Morse, Pharm.D.	State University of New York at Buffalo
Jeff Murray, M.D.	FDA
Maureen Myers, Ph.D.	Novirio Pharmaceuticals, Inc.
Thomas Myers, Ph.D.	University of Minnesota
Bach-Yen Nguyen, M.D.	Merck & Co., Inc.
Carla Pettinelli, M.D.	NIAID Division of AIDS
Richard Pollard, M.D.	University of Texas Medical Branch
John Pottage, M.D.	Vertex Pharmaceuticals, Inc.
Loulse Proulx, M.D.	BioChem Therapeutics, Inc.
Franck Rousseau, M.D.	Triangle Pharmaceuticals
Michael Saag, M.D.	University of Alabama at Birmingham
Renslow Sherer, M.D.	Cook County Public Health
Schlomo Staszewski, M.D.	Klinikum der J.W. Goethe Universitat
William Strain	AIDS Project Los Angeles
Michael Ussery, Ph.D.	Innogenetics, Inc.
Yenyun Wang, M.D.	Pharmacia & Upjohn, Inc.
Alastair Wood, M.D.	Vanderbilt University, Inc.