

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

**APPLICATION NUMBER: 20-154/S-028
20-155/S-020
20-156/S-021**

STATISTICAL REVIEW

Statistical Review and Evaluation

Labeling Supplement for:

Zerit (stavudine, d4T) NDA#20-412, 20-413

Videx (didanosine, ddI) NDA#20-154, 20-155, 20-156

Medical Reviewers:

Dr. Russ Fleischer

Dr. Joe Toerner

Background

The applicant, BMS, has submitted concurrent labeling supplements for their two products, Zerit (stavudine, d4T) and Videx (didanosine, ddI). This review covers both labeling supplements.

The labeling supplements are something of an anachronism. Both drugs were approved at a time when monotherapy studies formed the basis of approval for new HIV drugs. As a consequence, each drug had a monotherapy indication. Since then, the standard of care for HIV has become triple therapy, and new HIV drugs are studied in the context of multi-drug regimens. Indications are now written rather broadly: "for the treatment of HIV". The goal of the current labeling supplements is primarily to update the indication section to bring it up to date, and secondarily to provide additional information to clinicians regarding the drug use as part of a multi-drug regimen.

In support of this aim, BMS submitted the results of several trials which assessed the efficacy of d4T and/or ddI. Of these, two trials were identified by the FDA medical and statistical reviewers as most relevant. These trials, Start I and Start II, were the only large randomized studies with 48 weeks of comparative HIV-RNA data (the current standard for traditional approval of new HIV drugs) in which the drugs were used in triple therapy regimens.

Since the drugs are already approved, and since clinical data from these studies will not be going in the labels, the review of these studies is more limited than would be the case for the approval of a new HIV drug or for the addition of clinical data to the label. Rather, the aim of the review is simply to ensure that the known efficacy of the drugs when given as monotherapy also applies when given with other HIV drugs.

Start I

The medical review contains more detailed information about the design and patient population enrolled in this study. The study compared ZDV+3TC+IDV to d4T+3TC+IDV for 48 weeks, the primary endpoint was the percent of subjects with viral loads below the limit of the assay (1200 copies for ~~the~~ assay). A total of 204 subjects were randomized. Discontinuation rates were high but similar between the two treatment groups (36% and 35%, respectively). The primary analysis considered these patients to have HIV levels above the limit of the assay. At 48 weeks, the percent below 1200 copies was 54.4% on the control arm and 61.4% on the d4T arm. The 95% confidence interval for this difference was (-6.5% to 20.5%). The lower bound of this confidence interval, -6.5%, is well within the current range of equivalence for HIV studies at 48 weeks, which is plus or minus 10%. The DAVGs for CD4 counts during the 48 week study were +112 in the control arm and +150 in the d4T arm.

The primary concern regarding the efficacy results is the high dropout rate. However, there are four factors that mitigate this concern. First, the dropout rate was very similar between the two arms. Second, the applicant presented several methods of analyzing the results that accounted for missing data in different ways. The results appeared to be robust across these analyses. And third, the results of this trial are not being used to support a NDA or dosing change or support the inclusion of these data in the clinical trial section of a label. Therefore, the level of evidence need only support that the known efficacy of the drug d4T is not compromised when used in multi-drug settings. And finally, the results indicated a "strong" equivalence, that is, the d4T was numerically superior with a confidence interval that is not on the border of plus or minus 10%.

In conclusion, Start I does support the proposed d4T labeling change.

Start II

The medical review contains more detailed information about the design and patient population enrolled in this study. The study compared ZDV+3TC+IDV to d4T+ddI+IDV for 48 weeks, the primary endpoint was the percent of subjects with viral loads below the limit of the assay (1200 copies for _____ assay). A total of 205 subjects were randomized. Discontinuation rates were high but similar between the two treatment groups (37% and 44%, respectively). The primary analysis considered these patients to have HIV levels above the limit of the assay. At 48 weeks, the percent below 1200 copies was 48.5% on the control arm and 62.7% on the d4T arm. The 95% confidence interval for this difference was (0.1% to 27.7%). The lower bound of this confidence interval, -0.1%, is well within the current range of equivalence for HIV studies at 48 weeks, which is plus or minus 10%. The DAVGs for CD4 counts during the 48 week study were +111 in the control arm and +164 in the d4T arm.

There are two main concerns about interpreting the efficacy results from Start II. The first is the issue of dropouts that also applied to Start I. However, for the same reasons that applied in Start I (see above), the results do seem to support the limited efficacy claims being sought despite the high dropout rate.

The second issue is more problematic. Since the study design compares treatment arms with two differences (ddI+d4T compared to ZDV+3TC), it is not possible on the basis of this trial to distinguish the relative contributions of d4T or ddI alone. Had this study been part of an NDA application to support the approval of one or the other, it likely would be viewed as non-contributory. However, we already know that the drugs are effective on their own. Had this study used a non-standard control arm, again it would likely have been viewed as non-contributory. In this case, though, the control arm has been studied in its component parts, and it has been found that all three provide independent contributions to the efficacy of the regimen. Finally, had this study found a "weak" equivalence, where the ddI+d4T arm had the same or somewhat less proportion of successes at 48 weeks compared to control, the results may have difficult to interpret. In fact, the ddI+d4T arm was at least as good as the control arm and was arguably superior, although any claim of superiority would be suspect given the concerns raised above.

Start II therefore supports the proposed labeling change of d4T and ddI.

Michael Elashoff, Ph.D.
Division of Biostatistics III

Concur: Dr. Girich Aras