SGG

DEPARTMENT OF HEALTH AND HUMAN SERVICES

PUBLIC HEALTH SERVICE

FOOD AND DRUG ADMINISTRATION

PSYCHOPHARMACOLOGICAL DRUGS ADVISORY COMMITTEE

Thirty-Third Meeting

9:00 a.m.

Monday, November 19, 1990

Conference Rooms D & E
Parklawn Building
5600 Fishers Lane
Rockville, Maryland

MILLER REPORTING CO., INC.
507 C Street, N.E.

P A R T I C I P A N T S

Committee Members:

Daniel E. Casey, M.D., Chairman

John M. Davis, M.D.
Markku I. Linnoila, M.D.
Robert F. Prien, Ph.D.
Javier I. Escobar, M.Sc., M.D.
Jeffrey A. Lieberman, M.D.
Robert Mark Hammer, Ph.D.
Linda Frances Hezel, Ph.D.
Carol Ann Tamminga, M.D.

FDA Staff:

Paul Leber, M.D., Division Director
Thomas P. Laughren, M.D., Group Leader

C O N T E N T S

MILLER REPORTING CO., INC.
307 C Street, N.E.
Washington, D.C. 20002

P R O C E E D I N G S

DR. CASEY: Good morning. It is my pleasure to call this meeting to order and to welcome everyone to the 33rd meeting of the Psychopharmacological Drugs Advisory Committee. My name is Dr. Daniel Casey. I am from Portland, Oregon where I am on the staff of the VA Medical Center and the Oregon Health Sciences University. I have the pleasure of being the Chairman of this Committee.

Our first task should be to have other Committee members and people at the table introduce themselves.

DR. LEBER: I am Paul Leber, Director of the Division.

DR. PRIEN: Robert Prien, Director of Clinical Psychopharmacology, Division of Clinical Research of NIMH.

DR. HAMMER: I am Bob Hammer. I am a psychiatry and a statistics professor at the Medical College of Virginia.

DR. TAMMINGA: I am Carol Tamminga. I am professor of psychiatry at the University of Maryland.

DR. HEZEL: I am Linda Hezel and I am a professor of nursing at the University of Missouri, Kansas City.

MR. BERNSTEIN: I am Mike Bernstein. I am the Executive Secretary of this Committee.

DR. LIEBERMAN: Jeffrey Lieberman, Long Island Jewish Medical Center and Albert Einstein College of Medicine.

DR. DAVIS: John Davis, Illinois State Psychiatric

Institute.

DR. ESCOBAR: Javier Escobar, professor of psychiatry, University of Connecticut.

DR. LAUGHREN: I am Tom Laughren, Group Leader for Psychopharmacology at FDA.

DR. CASEY: Next, Mr. Bernstein has requested time for some administrative matters. Michael?

MR. BERNSTEIN: Thank you, Dr. Casey. I wish to welcome each of the Committee members to the 33rd meeting of the Psychopharmacological Drugs Advisory Committee. Additionally, I would like to welcome two of our new members to the Committee, Dr. Tamminga and Dr. Hezel.

My name is Michael Bernstein and I am the Executive Secretary of this Committee, which functions within the Division of Neuropharmacological Drug Products. Please bear with me while I make a few administrative announcements.

(Administrative announcements)

A review of the agenda by the Committee Management Branch personnel indicates that no Committee member requires limitations on their participation at today's session, based on reported interests as of November 15, 1990.

The following announcement addresses the issue of conflicts of interest and is made a part of the record to address even the appearance of such at this meeting:

It has been determined that all interests in firms

regulated by the Center for Drug Evaluation and Research which have been reported by the Committee members present no potential for an appearance of a conflict of interest at this meeting when evaluated against the scheduled agenda.

In the event that the discussions involve any products or firms not already on the agenda for which a special government employee has a financial interest, they are aware of the need to exclude themselves from such involvement and their exclusion will be noted in the record.

With respect to all other participants, we ask, in the interest of fairness, that they address any current or previous financial involvement with any firm whose products they may wish to comment upon.

Lastly, NDA 19-832, Sertraline, will be the only issue discussed by the Committee at this meeting. Thank you for your attention and this concludes my comments, Dr. Casey.

DR. CASEY: Thank you. Next we will have some opening remarks by Dr. Leber.

DR. LEBER: Good morning. My remarks are going to be very brief. I would like to welcome the Committee to Rockville once again to discuss this issue which obviously is of great interest to many people. I particularly want to extend a welcome to the two new members of the Committee, Dr. Tamminga and Dr. Hezel. Thank you and I hope you enjoy your association with us over the years.

Without further ado, I think we should get on with today's business, having gotten a few minutes late start. I would like to mention too that Dr. Markku Linnoila just came in and is now sitting on my right, at the left side of the table.

DR. CASEY: Welcome, Dr. Linnoila. It is a pleasure to have you here.

We will now move to the open public session, which is now in progress for those who would like to make a public comment. Is there anyone who wishes to address the Committee?

(No response)

I believe that no one has expressed interest in making a comment to the Committee at the open public session. We will, therefore, move on in our hearing.

The topic for today's Advisory Committee meeting is NDA 19-839, Sertraline safety and efficacy considerations. Dr. Hillary Lee, clinical reviewer for the Division of Neuropharmacological Drug Products, will be the first speaker. Dr. Lee?

PRESENTATION BY J. HILLARY LEE, Ph.D.

DR. LEE: As you all probably know, Sertraline is a new antidepressant. Animal studies suggest it is a potent and selective inhibitor of neuronal serotonin reuptake and has only very weak effects on norepinephrine and dopamine neuronal reuptake. In this respect it is similar to flu-

oxetine.

The parent compound has a terminal elimination half-life of 25 hours and the major metabolite, N-desmethyl-sertraline, which is believed to be far less active, has a terminal elimination half-life of roughly 60-100 hours.

The NDA submission contains 6 double-blind, placebo-controlled trials evaluating Sertraline in major depressive disorder. There were 4 outpatient trials, 2 of which showed Sertraline to be more effective than placebo and these will be the focus of my presentation today. A third outpatient trial compared Sertraline, amitriptyline and placebo and found amitriptyline to be more effective than placebo. The difference between Sertraline and placebo, while in the expected direction, was not significant.

The first study was a continuation trial and it supported the results of the positive trials. I will include some comments on it in my presentation. There were also 2 studies in inpatients which failed to demonstrate a difference between Sertraline and placebo.

(Slide)

The first study is protocol 103. This was a randomized, fixed-dose design, which compared 3 doses of Sertraline (50, 100 and 200 mg) daily and placebo, without titration of dosage.

The study began with a 4-day to 2-week single-blind

placebo washout which was followed by the 6-week double-blind phase. Patients were seen weekly for evaluations and 8 centers participated.

(Slide)

The subjects were outpatients with a DSM-III diagnosis of major depressive disorder and were required to have a minimum HAM-D total of 22 on the 17-item scale. Subjects were excluded if they required psychotropic medications beyond those allowed in the protocol, if they were not healthy or if they were resistant to antidepressants.

(Slide)

The primary efficacy assessments were the Hamilton Rating Scale for Depression, the Clinical Global Impression Scale and the Profile of Mood State. They yielded 5 key variables. The change from baseline for the HAM-D total score, HAM-D depression item, CGI severity item and POMS depression factor and the actual score for the CGI improvement item.

(Slide)

Some demographic characteristics: The mean age was 37-38 years but the range was 18-65 years. There was a preponderance of females in the Sertraline-treated group, not in the placebo. The majority of the patient diagnoses were major depressive disorder recurrent. That was approximately 60 percent of the subjects. That was followed by major

depressive disorder single episodes for 30 percent.

There was a total of 347 patients in the intent-to-treat population or approximately 90 patients per treatment group. The high dose Sertraline group had the smallest number of completers. Only 42 percent completed the trial.

(Slide)

Here are the reasons for dropouts. The first column is the number of patients entering the study. The final column is the total number that dropped out. The first two reasons, adverse events and lack of efficacy, were the only ones that differed among the treatments. High dose Sertraline had the most dropouts for adverse events, and placebo for lack of effect.

(Slide)

This slide on time of dropout shows that 63 percent of the high dose Sertraline dropouts occurred in the first 10 days of the trial (19 plus 14). This compares with 38 percent of the mid-dose group, and 29 percent of the placebo group, 28 percent of the low-dose Sertraline group. The effect of these early dropouts will be discussed quite extensively in the next section.

In the presentation of the results I will be focusing on two types of analyses which differ in the way they deal with the data from dropouts. The first is the last observation carried forward (LOCF) analysis, which is often

referred to as the endpoint analysis. In this analysis whenever a patient drops out of a trial, his scores at the time of dropout appear in each subsequent weekly analysis through the end of the trial. For example, if a trial is 6 weeks long and a patient dropped out at week 3, his week 3 scores would be used in the week 4 through week 6 analyses. The number of patients in the analysis is the same for each week.

The other type of analysis, the observed cases analysis, is often referred to as the completer analysis. In this analysis only subjects who are actually assessed at a time point are included in the analysis. If a patient drops at the beginning of week 4 after appearing for evaluation at weeks 1-3, he will be included only in the analyses for weeks 1-3. The number of subjects decreases from one week to the next when dropouts occur.

For each of the key variables, I am going to show you the results for the last observation carried forward and the observed cases analyses for the intent-to-treat population because the results of these analyses differ for a number of items. I would like to begin with the HAM-D depression item.

(Slide)

Within each table are change from baseline scores for each week and each dose. The asterisks indicate the significant 2-tail comparisons with placebo. The total group

is all 3 Sertraline groups combined.

In this table it is clear that the results for the LOCF and OC analyses were similar. The LOCF is above and the OC is below. Most of the comparisons from week 3 on were significant. Each dose was also significant, as were the scores for the combined Sertraline group.

(Slide)

The HAM-D total presents a very different picture. Here the 2 methods of analysis give different results. If we look at the end of the trial (week 6) the LOCF analyses were significant for 1/3 treatments, low-dose Sertraline, compared with 2 significant OC analyses.

(Slide)

The same difference between LOCF and OC analyses appears with the CGI severity score. At week 6 again, 1/3 treatments was significant for the LOCF analyses and 2/3 for the OC analyses.

(Slide)

For the CGI improvement score, 2 LOCF analyses and 2 OC analyses were significant at week 6.

(Slide)

In the POMS the majority of both the LOCF and OC analyses were significant from week 3 onward.

So when we look at the results for each variable, 2 variables were consistently significant by both methods, the

HAM-D depression item and the POMS depression factor. The other 3 variables showed weakness in the LOCF analyses.

(Slide)

When one looks at the results by individual treatment, there are several trends. This is the same 5 tables you saw before condensed into 1 slide. Here the asterisks indicate a significant comparison with placebo.

There are several trends I would like to point out here. As we saw in the individual slides, there were more OC than LOCF analyses significant for each treatment. These are the LOCF and these are the OC. As you can see, there is more down here for each treatment except the low-dose Sertraline group, which was the only treatment without numerous early dropouts.

Dropouts, particularly early dropouts, that occur differentially across treatment groups can affect LOCF analyses. So we see that high-dose Sertraline, with the largest number of early dropouts, had the smallest number of significant LOCF analyses. Among the individual treatments, the highest number of significant comparisons for both the LOCF and the OC occurred with the Sertraline combined group, the group down here.

The HAM-D total score had the smallest number of significant LOCF and OC comparisons. This is the one variable which includes a range of symptoms in this score.

When one examines the items on the HAM-D Scale, it is immediately apparent that a number of items reflect frequent side effects with Sertraline, for example, gastrointestinal symptoms, insomnia, agitation and anxiety. When these items are dropped from the analysis, the comparisons tend to be more significant.

One could also argue that the CGI items, by their global nature, have a range of symptoms, whereas, the 2 variables which focus narrowly on depressive mood are significant.

In conclusion, this study had problems using a fixed-dose design with immediate assignment to the maximum dose and probably would have fared better with titration to the maximum dose in each group. Patients were unable to adjust to 200 mg and to some extent 100 mg and large numbers dropped out early in the trial. This particularly affected the LOCF analyses where the high early scores are carried forward and bias the outcomes against the groups with early dropouts.

Protocol 104 is different from protocol 103 in a number of important ways. I will note these differences as we go along.

(Slide)

Study design -- one very important difference is flexible dosage. Notice also that in addition to placebo, we

now have a positive control. The duration of active treatment was 8 weeks instead of 6. Doses were titrated during the first 3 weeks and the remaining 5 weeks were considered the maintenance phase.

(Slide)

The entry criteria are similar, except for the HAM-D total score where failure to reach a minimum score of 18 on the first items or a decrease in total score of 25 percent or more during the baseline period was reason for exclusion.

(Slide)

Efficacy assessments are the same, except for the substitution of the symptom check list for the POMS.

(Slide)

Patient population -- the mean age is the same, 38-40 years and the range was 18-64. Sex ratio -- again, there is a preponderance of females, 54 percent in this study versus 53 percent in 103. The total population is somewhat larger, 448 subjects entered and there are 427 in the intent-to-treat population.

Completers -- another big difference. In this trial the number was much higher, 60 percent against 52 percent. In addition, the proportion of completers was similar among treatments.

(Slide)

Dropouts -- 40 percent versus 48 percent in the

fixed-dose trial. The reasons for dropouts are primarily adverse effects for active treatment and lack of effect for placebo.

(Slide)

The maintenance dosage -- 104 for amitriptyline. Notice that the average for Sertraline, 145 mg, was not permitted by the protocol. The doses were 50, 100 and 200 for Sertraline.

(Slide)

These are the efficacy results for protocol 104 in a summarized form. The population used was the intent-to-treat group. The results are given for 2 time points, week 5, the time when 70 percent of each group were still present, and week 8, the final week of the trial, and again for the LOCF and OC analyses. The change from baseline is shown for each of the 5 variables. An asterisk indicates a significant comparison with placebo. All comparisons for Sertraline and for amitriptyline were significant for the first 4 variables. With the SCL depression factor, the 3 amitriptyline comparisons were significant and 2 of the Sertraline, although at a less significant level. These results are supported by the analyses done for other time points which generally indicate a significant superiority over placebo from week 3 onward for both Sertraline and amitriptyline.

(Slide)

The results for the HAM-D total score are also shown graphically. In the LOCF analyses the scores drop from approximately 23 at baseline to 15 at the end of the trial for placebo, the least improvement, to 12 for Sertraline and to 11 for amitriptyline, the most improvement.

(Slide)

The same is true for the OC analyses. The final placebo score is 12; Sertraline, 8; amitriptyline 7.

In conclusion, this study provided consistent evidence that Sertraline produced more improvement than placebo and this was the case for all variables.

I would also like briefly to mention a 1-year trial, protocol 320. This trial treated 467 depressed patients with open Sertraline for 8 weeks. Those patients who had a satisfactory response were then randomized to double-blind treatment of Sertraline or placebo for 44 weeks. So it was an open trial for 8 weeks and double-blind for 44 weeks.

The idea of including maintenance studies in an NDA for depression is one which we, at the FDA, would like to foster because such studies more closely approximate the way antidepressants are used.

Protocol 320, however, had major problems. In the protocol the critical variables were either not defined at all or the definition was insufficiently detailed to be

useful. For example, in the first critical decision, namely, to decide whether or not the patient had shown a satisfactory response to open Sertraline, an objective definition of satisfactory response was not provided.

Another critical variable that was not defined was what constituted a relapse in the double-blind phase. Relapse is a major outcome variable in maintenance studies.

Because the protocol did not describe these major decisions adequately, they became post hoc decisions with the corresponding problem of data-conditioned analyses. There were other problems, including a very large number of investigators (39) and various protocol violations. Also over 60 percent of the patients in both groups received concomitant psychotropic medications.

Dropouts because of inadequate therapeutic response were much higher in the placebo group. I am just going to talk about one finding that interested me because the investigators checked on the case report form why they were dropping the patient. So I looked at the inadequate therapeutic response dropouts. Dropouts for this reason were much higher in the placebo group (42 percent) than in the Sertraline group (9 percent). Of all placebo dropouts for lack of treatment effect, 48 percent occurred in the first 8 weeks of the double-blind phase. Only 29 percent of Sertraline dropouts for lack of treatment effect occurred in this same

period.

Because of the serious protocol deficiencies, results of this study cannot be considered definitive. The dropout data do suggest, however, that once Sertraline has been started, it may be advisable to continue it beyond 8 weeks.

My conclusion about the efficacy section of the NDA is that the sponsor has provided two studies, one of which is more consistent than the other, to demonstrate the effectiveness of Sertraline.

DR. CASEY: Thank you, Dr. Lee. In addition to our presentations scheduled today, next will be Dr. Nevius, from the Agency.

DR. LEBER: For the record, Dr. Nevius is a Group Leader in the Division of Biometrics.

PRESENTATION BY ED NEVIUS, Ph.D.

DR. NEVIUS: Thank you. I would just like to highlight a few of the elements of the statistical review of an NDA that are pertinent to the Sertraline NDA that we are discussing now.

(Transparency)

The points I will briefly touch on are shown on the overhead. First, is the design of the studies appropriate for the objectives?

With regard to this first point, Dr. Lee has

already pointed out problems with regard to the design of two of the studies, namely, the absence of titration in study 103 which caused a large number of early dropouts, particularly in the 200 mg Sertraline group, and the design problems in study 320 with regard to the inadequate protocol definition of endpoints.

(Transparency)

With regard to the data set analyzed, we always like to see an analysis utilizing all patients randomized to the study in addition to other various evaluable patient type analyses which may be performed by the sponsor. We call this the intent-to-treat population of patients and usually define it to be all patients randomized to the study who actually received at least one dose of the assigned treatment. Analyses of this population will avoid possible biases caused by patient deletions made after randomization for various reasons.

(Transparency)

With regard to the statistical analyses of the data, we check whether we agree with the specific analyses and methods for performing those analyses, as well as whether other equally plausible analyses produce reasonably similar results.

For example, in study 103 which, you remember, involved three dose levels and placebo, the sponsor relied

upon an analysis which combined data from these three separately randomized groups and treated them as one group. Dr. Choudhury, the statistical reviewer for this application, performed a test for dose response to see if there was evidence of increasing efficacy with increasing dose. Dr. Choudhury will show the results of a non-parametric test for this purpose, called the Jonckheere-Terpstra test for ordered alternatives.

(Transparency)

In some drug classes where the primary endpoint is survival it may make sense to talk about the intent-to-treat analysis. In that case, patients who do not remain in the study until the end may be treated by the statistical analysis as censored observations and their time in study will still be taken into consideration in the analysis.

In trials of depression this situation is more complicated. If the patient is not rated in the study, then there is no optimal way to include the patient in the analysis. The LOCF (last observation carried forward) analysis, often called endpoint analysis when restricted to the last time point in the study, has been widely used in psychopharmacological drug trials but has inherent biases when there are differential dropout rates. Patients generally get better over time, that is, their HAM-D scores go down.

As pointed out by Dr. Lee, the LOCF analysis, for

example in study 103, will then carry forward a larger proportion of early high HAM-D scores for the high-dose group than the placebo group, making it difficult to show efficacy for the high-dose group by this analysis.

Incidentally, this problem also arises in attempts to show evidence of dose response. You will see in a minute that Dr. Choudhury has also analyzed study 103 for dose response with this high dose group with large early dropouts (200 mg group) omitted. On the other hand, the observed cases analysis completely ignores the effect of a large proportion of randomized patients being left out of the analysis. So we look for similar results for these 2 types of analyses and look for reasons for any discrepancies. Clearly, methodological research is really needed in this area in the future.

(Transparency)

Finally the last two points, we look at how similarly various standards behave in a multicenter trial, as well as how results of the various studies in the NDA compare. Dr. Choudhury will present graphs showing results by center in studies 103 and 104, as well as descriptive statistics comparing the various studies in the NDA.

We should be aware that multicenter studies were designed to be able to show significant results from the primary combined analysis and nominally significant results

for a center analyzed alone are not usually expected. But, on the other hand, nominally significant results for a few out of many centers in one multicenter trial should not routinely be taken as independent corroboration of efficacy.

In conclusion, I would like to mention that the p values we calculate are only approximations as various assumptions for statistical analyses never exactly hold and there is rarely one correct p value for making conclusions about a particular study. While dropouts are problematic in the efficacy studies for this NDA, I believe it is clear that study 104 produces significant results no matter how the data are analyzed. Study 103 has similarly robust results for 2 variables, including the HAM-D depressed mood item.

Dr. Choudhury will now present some statistical information about the two studies. A few of the tables will contain more information than you can absorb by a brief look at the slides but we have provided copies of his tables and his graphs in your handouts for later reference. You may notice one slight discrepancy between the p values that we have been quoting and that Dr. Choudhury will quote and the ones that I believe the sponsor will quote. The FDA presentation gives 2-sided p values for Sertraline versus placebo comparisons, while the sponsor's slides I believe will give 1-sided p values for these comparisons. Dr. Choudhury will show some of his graphs and tables for the primary results

now.

DR. CASEY: Thank you, Dr. Nevius. Next will be the presentation by Dr. Choudhury, who will give us a statistical review.

PRESENTATION BY JAPOBRATA CHOUDHURY, Ph.D.

DR. CHOUDHURY: Good morning. The various graphs and tables I shall be presenting are contained in the handout supplied to the Committee. First I will go over study 103. I would like to show you graphically how the dropouts occurred over time in study 103.

(Slide)

This graph shows the retention rates by treatment groups over time. Similar graphs, giving the percentages in parentheses, are in the handouts. The light green is for the 200 mg dose. The early losses in the 200 mg group are evident with only 43 percent of the patients in this group completing the study. The high dropout rates and differential dropout rates among the treatment groups lead us to look at both OC and LOCF results for various analyses.

(Slide)

First we will look at an analysis to investigate whether a dose-response relationship exists among the treatment groups. The Jonckheere-Terpstra test is a non-parametric test for this purpose. All these p values are highly significant and the HAM-D item 1 clearly shows

evidence of positive dose response, while the HAM-D and CGI improvement show at least marginal evidence of dose response at week 6 when the 200 mg dose is excluded.

(Slide)

This slide and the next one gives results for the pair-wise comparisons between each dose group and placebo. These results are for the HAM-D totals which were rather inconsistent across different analyses. The orange color is for LOCF and the green color is for OC.

(Slide)

HAM-D items, on the other hand, show quite robust results favoring Sertraline.

(Slide)

This slide and the next one gives results by center. This one is for HAM-D total. Results for larger centers (the first 5) show consistent numerical superiority for both 50 mg and 100 mg doses over placebo, although the relationship between 50 and 100 mg is not clear-cut.

(Slide)

The same is true for HAM-D item 1.

(Slide)

Now I will go over study 104. This slide of retention rates shows more closely comparable retention rates for the 3 treatment groups than was the case for study 103. Approximately 60 percent of the patients completed the trial.

Although the dropout rates are more similar among treatment groups than in study 103, we still want to look at both OC and LOCF results due to the rather high overall dropout rate.

(Slide)

This slide and the next give results for the pairwise comparisons between each active treatment group and placebo. These are the results for the HAM-D total. All the values are highly significant, although results for amitriptyline are more significant. The ordering of placebo showing less improvement than Sertraline and Sertraline showing less improvement than amitriptyline is clear.

(Slide)

The same comment applies to HAM-D item 1. It can be noticed that in study 104 the p values for HAM-D totals are more significant than those for HAM-D item 1. The reverse is true for study 103.

(Slide)

This slide shows 95 percent confidence intervals for the Sertraline-placebo comparison by center for HAM-D total at last visit. All centers show numerical superiority for Sertraline with centers 2 and 14, suggesting evidence of efficacy individually, as indicated by the confidence intervals excluding, or nearly excluding zero.

I am not going to talk about study 320 due to time [illegible] but there are several pages in the handout for

your reading.

(Slide)

Finally, this is comparison of the studies. Detailed comparisons of all studies is not possible here. Let us compare the studies with respect to HAM-D total first at week 4. Study 320 is of a different kind to investigate the efficacy of Sertraline in the prevention of relapse and is not considered here. Studies 104 and 315 did not have separate dose groups for Sertraline. From other studies we picked up only the 100 mg group for a clear comparison graphically. These latter studies did not have an amitriptyline group. Except for study 310, treatment differences appear reasonably consistent over the studies. Study 103 has the highest treatment difference even though the p values are not quite significant, probably because of a smaller number of patients per treatment group compared with study 104. If all the dose groups of Sertraline are combined, study 103 results become highly significant.

(Slide)

Comparison at week 6 -- only 3/5 studies considered in the last slide had durations of 6 weeks or more. We still see 5 columns because we included LOCF results for studies 103 and 104. At week 6, the treatment difference for study 104 was no bigger than for other studies but, because of larger numbers of patients per treatment group, study 104

produced highly significant p values. Results for amitriptyline are consistently numerically superior to those for Sertraline. Negative values indicate improvements.

(Slide)

Conclusions -- study 104 provides highly significant p values by almost all analyses. Study 103 provides significant p values with respect to HAM-D impression item and POMS depression-factor. Despite the fact that the protocol for study 320 is not highly satisfactory, study 320 provides some sort of supporting evidence for the short-term efficacy of Sertraline in prevention of relapse of depression.

Descriptive statistics show numerical superiority of Sertraline consistently over all studies at most time points.

DR. CASEY: Thank you, Dr. Choudhury. We will next move on to a presentation by Dr. Knudson on the safety data.

PRESENTATION BY JAMES F. KNUDSON, M.D., Ph.D.

DR. KNUDSON: One of the jobs of the clinical review is to ensure that new drugs are reasonably safe and effective before they are marketed. Dr. Lee has discussed the latter aspects and I will discuss the safety issues. Together these presentations, coupled with the input from other sources today, should enable a risk-benefit assessment of Sertraline.

(Slide)

There are 4908 subjects in the Sertraline data base. The majority of the active control drug was amitriptyline. Of the Sertraline-treated population, 1902 are the depression studies; about a third of the number on obesity and, as you can see, only a handful on panic. A total of 861 Sertraline-treated patients, 372 from the depression and another 489 in obesity, participated on placebo-controlled dose titration studies.

The safety data presented today will utilize this pool of patients since it is our view that all patients exposed to a drug, whatever the reason for treatment, can provide relevant safety data.

(Slide)

Of the Sertraline-treated patients, 3/5 (58 percent) were female; 38 percent were non-white and 1/6 was 65 years of age or older. The figures for placebo and active are comparable.

(Slide)

Twenty-two percent of patients received drug for more than 90 days and 3 percent for longer than a year.

(Slide)

On the question of discontinuation in placebo-controlled studies, 3 groups, as you can see from this slide, were comparable, with about 2/5 subjects leaving prior to completion of the study, however, as expected, mostly for

adverse events on the active and for lack of efficacy on the placebo.

(Slide)

In our safety assessments here we look at major events, such as deaths and dropouts across all studies, first to provide an overall, albeit rather coarse, picture of the drug. Seven deaths occurred, 4 on Sertraline, 1 on active and 2 on placebo. In each case, the treating physician stated that the study drug was not responsible for the death.

Having reviewed the supporting documents, it is my view that none of the deaths can be reasonably attributed to Sertraline use.

(Slide)

Realizing the difficulty in interpreting data where analyses ignore differential exposure time, this table does show that disproportionate numbers of suicide attempts do not occur among the 3 treatment groups. All suicide attempts appeared in depressed patients, none in the obese. As noted in the last slide, for 2/9 Sertraline-treated patients who committed suicide, the suicides were complete.

Suicidal thinking was measured by the Hamilton Psychiatric Rating Scale for Depression, HAM-D 3. Emergence of serious suicidal thinking is defined as a change from a HAM-D score of 0 or absent or 1 (self-reproach) to a score of 3 where ideas or gestures of suicide occur, to 4 where

attempts are actually made. Very few patients got this much worse on treatment. For the final time point or OC analysis, the incidence of this event was very low in all treatment groups, no more than 1 patient in any treatment group had this event.

For the last patient visit, or LOCF analysis, the incidence rates were uniformly greater than the final time point analysis. The incidence rate in the placebo group, as you can see in protocol 104, of 6 percent certainly was higher than the incidence rates for the other 2 treatment groups which were of comparable magnitude.

In the Sertraline development program there were 3 non-fatal cases of active overdosage, either with Sertraline alone or in combination with other drugs. Doses of up to 210 mg were reported. The patients recovered and no specific therapy was required.

(Slide)

This slide provides the side effect profile of Sertraline in the placebo-controlled titration studies. Events listed were reported by at least 1 percent of the Sertraline-treated patients and were statistically significant compared with the placebo control.

Treatment-emergent side effects and dropouts occurred most frequently in 4 body systems, namely, the gastrointestinal, the psychiatric, the central nervous system

and peripheral nervous system and the autonomic nervous system.

As would be expected with a serotonin reuptake blocker, the most commonly reported adverse experiences and discontinuations were GI, 54 percent for the Sertraline; 26 percent for the placebo and 26 percent for the active.

With respect to sexual dysfunction among Sertraline-treated patients, this occurred primarily in males and at statistically significantly greater rate compared with the placebo and active. The preponderance of complaints was related to transient ejaculatory disturbances. Complaints of impotence and loss of libido occurred less frequently. Few patients discontinued for this event.

Patients exposed to Sertraline also reported insomnia more frequently as a side effect compared to the placebo and active control-treated patients. Again, few patients discontinued.

Somnolence, which incorporates the terms sedation and drowsiness, was reported more frequently in the Sertraline-treated patients versus the placebo-treated and less than the active control group. Sedation is, of course, one of the most common side effects of TCAs.

Dizziness, twitching and tremor were the 3 most common CNS/PNS adverse events in Sertraline-treated patients. One percent of Sertraline- and placebo-treated patients,

compared with 4 percent of the active control drug group discontinued due to dizziness. No patient discontinued due to tremor. Twitching was mild to moderate in intensity and was reported by all treatment groups. No patient discontinued.

As you can see, CNS/PNS side effects occurred overall more frequently in the active control group and 69 percent of the active control group reported dry mouth compared to 16 percent of Sertraline and 9 percent of the placebo. One percent of the Sertraline patients discontinued due to dry mouth compared to less than 1 percent of placebo and 5 percent active controls. The symptom of increased sweating, for the most part, was mild and no patient discontinued.

(Slide)

This table displays a list of clinical events which occurred significantly more often in active than with placebo. Increased appetite and weight gain, side effects associated with TCA use, appear less frequently in the Sertraline-treated patients. Indeed, an interesting consequence of Sertraline use has been a clinically significant weight loss, a decrease of greater than or equal to 7 percent below baseline, in some Sertraline-treated patients compared with both placebo and active.

(Slide)

There are 2 other clinically important events observed during development, namely skin rash and purpura. This table displays the incidence of skin rash in placebo-controlled titration studies. There are no differences. Rash and possible allergic events have been reported with other Sertraline reuptake inhibitors. Six Sertraline, 1 placebo and 1 active discontinued due to skin rash.

(Slide)

Synopses of the individual cases in the depression studies where a rash was reported reveal the following: All but 1 were female; age was equally distributed amongst the young and old; 4/6 had a positive history for allergies. Rash, for the most part, was reported to be rather moderate or moderate to severe in intensity. Adjunctive therapy was used in all but 1 case. In 5/6 cases there were no associated signs or symptoms -- this is important -- related to fever, respiratory symptoms, lymphadenopathy, arthralgias and abnormal laboratories.

In the last case there was a report of facial edema and joint swelling, accompanied by leukocytosis and eosinophilia. The patient was hospitalized, treated with dexamethasone, triphinedine (phonetic) and betamethasone ointment and recovered. The final diagnosis was photosensitivity to Sertraline.

An additional patient in the obesity study discon-

tinued due to severe erythematous maculopapular rash, which was characterized by the consultant dermatologist as erythema multiforme. This patient had also received trimethoprim sulfamethoxasole 5 days prior to the rash. Consequently, the relationship of the event to the Sertraline exposure remains unknown.

(Slide)

The incidence of purpura in placebo-controlled titration studies is displayed here. You see that whereas both Sertraline- and placebo-treated patients reported purpura, there were no reports in the active control group. One patient exposed to Sertraline discontinued the study due to a recurrence of what was considered to be DTP.

(Slide)

This slide lists the patients in the placebo-controlled titration studies who had reports of purpura. The patient I just alluded to, patient 0045, as you can see, was the only patient who had a platelet count lower than 150,000. However, the patient had a low baseline count of about 130,000 which was unchanged during the trial.

So in review of the individual cases, purpura appeared to be an isolated event. There was no evidence of thrombocytopenia and, once again, purpura has been reported with other serotonin reuptake inhibitors.

In light of the fact that activation of hypermania

has been reported in patients treated with antidepressants, the occurrence of this event was assessed. In placebo-controlled depression studies, hypermania or mania occurred infrequently and there were no differences between the incidence rates among the three treatment groups. One Sertraline-treated patient in these studies discontinued.

Overall, in the Sertraline development program, 10/2,710 patients (about 0.4 percent) either had a dose reduction or discontinued as a result of mania.

Except for liver function tests, the incidence of clinically significant laboratory abnormalities in the Sertraline development program was low.

(Slide)

This slide displays the incidence of elevated LFTs in all placebo-controlled studies. The criterion is 3 times the upper limit of normal. There was no statistically significant difference between the 3 groups.

In all cases examined, all patients were asymptomatic and there were no reports of cholestasis among the patients. Values were found to return to baseline after treatment was discontinued in all cases. There were no statistically significant differences between any of the treatment groups with respect to the number of patients who discontinued due to elevated LFTs.

In addition to the few reports of elevated LFTs,

statistically significant reductions from baseline serum uric acid levels occurred in the Sertraline-treated patients compared with placebo and active control. The mean change from baseline was approximately 7 percent. The clinical significance remains somewhat unclear, although hyperuricemia has been reported in patients treated with triazolam benzodiazepines and with thioxanthene. Additionally, and importantly, hyperuricemia is often associated with hyponatremia related to SIDH.

Other safety data examined were vital signs and ECGs. With respect to the former, Sertraline use was not associated with any clinically meaningful changes in vital signs. Overall in the clinical development program, the incidence of patients with any ECG marked as more abnormal was similar in the Sertraline and placebo groups.

(Slide)

The last table displays a line listing of individual cases of patients who discontinued due to ECG abnormalities. The rate of dropout was not different across the drug groups in all the multiple dose studies.

Several studies have been implemented by the sponsor to assess the potential for drug interaction with Sertraline. Two finalized reports have discussed the possible interaction between Sertraline and tolbutamide and Sertraline and lithium. A statistically significant decrease

of 15 percent in tolbutamide clearance was reported by the sponsor and seen in patients treated with Sertraline compared with placebo. No statistically significant differences between Sertraline and placebo were seen in the renal clearance of lithium.

Lastly, as you recall from the earlier slides, a fair number of Sertraline-treated patients in the development program were 65 years of age or older. A comparison of the incidence of side effects in elderly with younger patients in the depression studies revealed that the proportion of Sertraline patients experiencing side effects was similar in the 2 age groups, as was the frequency of occurrence of specific side effects. Additionally, the incidence of laboratory abnormalities was generally similar.

Overall, based on the safety data, it is my view that the sponsor provided evidence that Sertraline is safe when used in the treatment of depression.

DR. CASEY: Thank you, Dr. Knudson. Do any of the Committee members have questions that they would like to address to the presenters for clarification or amplification?

(No response)

I guess not. That being the case, we have the opportunity to stretch our legs and have a short coffee break. We will come back in 15 minutes, at 10:35. Thank you.

(Brief recess)

DR. CASEY: I hope everyone feels fresh and ready to go on to part two. The first thing I would like to do is to thank and compliment the Agency's presenters on their clear and focused presentations. They were very well done, very easy to follow.

The next section will be to move to the sponsor's presentation by Pfizer Pharmaceuticals. That will be presented by Dr. Steven Ryder.

PRESENTATION BY STEVEN RYDER. M.D.

DR. RYDER: Good morning. Thank you, Dr. Casey, Dr. Leber, Dr. Laughren. I would first like to express our appreciation of and agreement with the comprehensive reviews of Drs. Lee, Choudhury and Knudson. During the past few months, the entire FDA review team, including Dr. Laughren, has been very helpful in assisting our preparation for this meeting.

I will try to minimize repeating areas that have been well covered by the prior speakers. I am going to not cover some areas and quickly pass through some others. If you would like me to cover any of these areas in greater detail, please just let me know.

(Slide)

The outline of my presentation is that I will cover just one or two items regarding efficacy, clinical pharma-

cology and, finally, safety. I think that the review presented for protocols 103 and 104 has been comprehensive and I am not going to make any comments on the review.

(Slide)

I would like to mention some things about protocol 320. This study, as you have heard, was a large, multicenter study where patients with depression were treated for 8 weeks with open-label Sertraline, using an upward dose titration regimen from 50-200 mg/day. After 8 weeks, those patients judged to have a satisfactory clinical response on the basis of investigator opinion were offered double-blind randomization. Those patients entering into this double-blind period were randomized for 44 weeks of double-blind treatment with either Sertraline or placebo. The dosage was started low and the protocol called for increases up to 4 capsules per day, corresponding to 200 mg/day of Sertraline, for worsening depression. If this dose was ineffective, the patient was to be discontinued. An entrance criterion prior to entering the initial 8-week open-label period was that the patient had a minimum HAM-D total score of 17.

(Slide)

There were 467 patients who entered the initial open-label treatment period and 350 patients completed. Of these 350, 295 accepted randomization into the double-blind treatment period. The randomization scheme called for a 2:1

ratic. There were, therefore, 185 Sertraline patients and 110 placebo patients.

I think it is important to note that at this point of randomization the groups seemed balanced in terms of demographics, disease characteristics and disease severity. Backup slides are available if you would like to see this.

As was mentioned, there was a difference in the percentage of patients who completed the double-blind treatment period, with 58 percent of the Sertraline and 33 percent of the placebo patients completing this period. This difference was principally due to a difference in the percentage of patients discontinued due to lack of efficacy, with 41 percent of the placebo patients and 8 percent of the Sertraline patients discontinuing for that reason.

(Slide)

As mentioned, efficacy analysis for this study was different from the other trials. A principal efficacy variable was the relapse analysis. This slide presents the results of the intention-to-treat, last observation carried forward, for the relapse analysis. The definition of relapse is on the last two lines of the slide. Relapse was defined as becoming and remaining ill where illness was defined as having a CGI severity score of greater than three. As you know, this is a seven-point scale. So that would be the higher end of the scale. There was a difference in the

percentage of Sertraline and placebo patients that had relapse using their LOCF data.

(Slide)

Another way to review efficacy in this trial is to examine the discontinuations due to lack of efficacy. This figure presents the Kaplan-Meier curves using the endpoints of discontinuation due to lack of efficacy. The curves are presented for the Sertraline group in yellow, the placebo group in white, and the boundaries of the 95 percent confidence intervals are also shown. There was a highly statistically significant difference in the curves of the Sertraline and placebo groups.

(Slide)

Several other Kaplan-Meier analyses were done in conjunction with FDA and this is another. This is a Kaplan-Meier curve using the endpoint of CGI severity score greater than 3. That is, the first occurrence of first illness, accepting the definition of illness as a CGI score of 4 or more.

Once more, the curves for the Sertraline and placebo groups are shown with the limits of the 95 percent confidence intervals. Once more, there is a significant difference between the groups.

(Slide)

This slide presents the efficacy conclusions and

they are in agreement with those that you saw previously. That is, based on the results of protocols 103 and 104, Sertraline, within a dose range of 50-200 mg/day, is more effective than placebo in the treatment of major depression.

Second, based on the results of protocol 104, Sertraline causes improvement similar to amitriptyline.

Third, based on the results of protocol 320, the results of which were just discussed, following successful treatment with Sertraline, continued Sertraline therapy reduces the incidence of depressive relapse compared with placebo.

I will now turn to clinical pharmacology. The next two slides list the clinical pharmacology studies and drug interaction studies in the Sertraline development program.

(Slide)

This is a list of the clinical pharmacology studies. Several studies, such as the kinetics in renal insufficiency, were recently completed and the data are being analyzed. Others, such as the kinetics in hepatic insufficiency, are ongoing. I will not discuss these studies.

(Slide)

This slide lists the drug interaction studies. I would like to mention that an attempt was made to examine drugs that would be representative of classes commonly co-prescribed with Sertraline, as well s classes with different

metabolic pathways. Once more, some studies have data being analyzed and will not be discussed.

(Slide)

Following oral administration of Sertraline, peak plasma concentrations (C_{max}) was typically reached 7-8 hours following oral ingestion (T_{max}). Linear pharmacokinetics were demonstrated in a dosage range of 50-200 mg. The terminal elimination phase was achieved 12-16 hours following oral ingestion. The average terminal elimination half-life from plasma was approximately 24-26 hours.

I should mention here that the Sertraline plasma concentration achieved after administration of Sertraline for 14 days once daily was consistent with the 24-26-hour half-life that you see here and was dose proportional.

I should also mention that when Sertraline was administered with food, there was a 32 percent increase in C_{max} and a 39 percent increase in AUC. Accordingly, current recommendations are that Sertraline be administered with food.

(Slide)

Sertraline is extensively metabolized and an important initial pathway of metabolism is to N-desmethyl-sertraline. N-desmethylsertraline, as Dr. Lee mentioned, has a plasma terminal elimination half-life from 62-104 hours. Importantly, both biochemical and pharmacological studies

show that N-desmethylsertraline is substantially less active than sertraline. Specifically, N-desmethylsertraline was approximately 1/10 as potent as the parent compound in *in vitro* activity of serotonin reuptake inhibition. N-desmethylsertraline was inactive in the mouse Porsolt model of depression and displayed innocuous behavioral effects at doses up to 1 gm/kg in mice and rats.

(Slide)

I mentioned that a number of Sertraline drug interaction studies have been done. The general design for these studies was that the co-administered drug was given before and after three weeks of either Sertraline or placebo administration. Three weeks was chosen as being more than adequate in terms of allowing Sertraline or desmethylsertraline to achieve steady state.

With warfarin there was an 8.9 percent increase in prothrombin time after Sertraline compared with placebo. Although this change seems small, prudent medical practice would likely suggest that prothrombin time be monitored when Sertraline is started or stopped.

Compared to placebo, Sertraline caused a 6.5 percent increase in steady-state serum lithium levels and a 1.9 percent decrease in lithium clearance. Neither of these changes was statistically significantly different from placebo.

I should mention that there is little clinical experience with the co-administration of Sertraline and Lithium. As Dr. Lee mentioned, a few patients entered our study with the diagnosis of bipolar disorder. There was a 16 percent decrease in tolbutamide clearance and a 13 percent decrease in diazepam clearance. Both of these decreases are considered to be of minimal, if any, clinical importance.

Sertraline has no effect on beta blocker pharmacodynamics. The specific beta blocker used in this trial was atenolol. That is the CD_{25} of isoproterenol, the dose that increases heart rate. The chronotropic dose of 25 beats/minute was unchanged compared with placebo. Finally, the pharmacokinetics of digoxin were unchanged by Sertraline.

(Slide)

This slide presents the approaches that were taken to examine the Sertraline data base concerning safety issues. I would like to say that we are in complete agreement with Dr. Knudson's review. I think it was extremely comprehensive. I would only like to slide to the very last item, ECGs, and I would like to note that a special effort was made, and a further and more detailed review of all ECG tracings from U.S. and Canadian depression trials was done in a blinded manner by Dr. Charles Fisch and Dr. Suzanne Knoebel, of Indianapolis. Dr. Fisch is here today and he is prepared to answer any questions regarding this analysis.

The analysis included change from baseline to final visit in heart rate and the ECG intervals listed, as well as a review of the incidence of conversions from baseline to final visit for overall ECG interpretation, rhythm, axis and the other listed parameters.

(Slide)

There were no significant changes in the ECG intervals of Sertraline and placebo patients. As expected, amitriptyline resulted in an increase in heart rate and a consequent reduction in the ECG RR and QT interval.

(Slide)

This shows the number of conversions from normal at baseline to abnormal at final tracing in specific categories of ECG changes. There were no differences between Sertraline and placebo. The statistically significant difference in the amitriptyline group was because -- and Dr. Fisch can explain this further if I am not clear -- amitriptyline, by accelerating the heart rate, to a great extent precludes recurrence of sinus bradycardia which is a normal occurrence. You can see that its frequency in the Sertraline and placebo groups is the same.

The review by Drs. Fisch and Knoebel concludes that Sertraline has no demonstrable effect on the clinical electrocardiogram.

(Slide)

The overall conclusions are that Sertraline, in the treatment of major depression, is effective. It appears to reduce the incidence of depressive relapse. Sertraline offers well-tolerated therapy and is without significant anticholinergic CNS arousal, cardiotoxic activity and does not cause weight gain.

Sertraline appears to offer effective and well-tolerated therapy in the treatment of patients with major depression.

Thank you and we will be glad to answer any questions.

DR. CASEY: Thank you, Dr. Ryder. Are there any questions or comments from the Committee? Dr. Prien?

DR. PRIEN: In terms of clinical results from studies 103 and 104, we have been provided with mean change scores from baseline and the outcome of statistical analyses for the difference in mean change scores between drug and placebo. However, it is difficult to determine the clinical significance of the statistically significant differences in mean change scores, particularly when you are looking at differences in change of only 2 or 3 points on the HAM-D total score or differences between groups on the CGI improvement and severity of illness scale of 0.4 or 0.5, which is not a tremendous difference.

Do you have data on the percentage of patients in

each group who showed improvement based on some designated criteria-based outcome measure? Say, CGI improvement score of 1 or 2 or HAM-D total score of less than 10, or even percent reduction over baseline for individual patients?

DR. RYDER: One of the analyses done was an examination of the proportion of patients who responded. Two definitions of response were looked at. One was a HAM-D decrease of greater than 50 percent from baseline. A second was, as Dr. Prien suggested, a CGI severity score of 1 or 2.

These analyses showed that in protocol 104 the proportion of patients improved was generally comparable between Sertraline and amitriptyline. If I remember correctly, the figure is about 65 percent using intention-to-treat, last observation carried forward.

DR. PRIEN: I am somewhat less impressed with comparing Sertraline with amitriptyline, particularly since your mean dose was about 100 mg of amitriptyline, which, for patients who have HAM-D scores of over 18, may be somewhat small. But I am interested in the comparison between Sertraline and placebo with respect to individual improvement rates.

DR. RYDER: If I remember correctly, the correspondent placebo proportion was about 45, 48 percent.

PFIZER REPRESENTATIVE: Forty-eight.

DR. RYDER: I am told it was 48 percent.

DR. ESCOBAR: Is that only for 104 or also for 103?

DR. RYDER: For 103 we also have proportions for the combined Sertraline group, as well as the individual Sertraline groups. If my memory serves me correctly, the placebo proportion is about 45 percent for protocol 103 and for the Sertraline combined group it is approximately 58 percent. I am sorry, 60 percent. There are people who have the data available on the side.

(Slide)

This is the evaluable patients subset, with the proportion of patients going from week 1 to week 8. The proportions there are presented for Sertraline, amitriptyline and placebo.

DR. LINNOILA: In the package provided to us we had studies which were not discussed today at all. Reading that package, I come to the conclusion that Sertraline is efficacious basically for outpatient treatment but I do not see any efficacy in the inpatient trials. Is that a fair reading of the data or am I somehow mistaken?

DR. RYDER: There were two inpatient studies. I believe Dr. Lee mentioned them initially. These were studies that had from 23-37 patients per group. They were 4 weeks in duration and they did not show a difference between Sertraline and placebo.

I should mention that in the preparation of our

studies for presentation, we have looked at our data base in protocol 104, looking at patients who had a HAM-D baseline score from 17-24 versus those who had a score of 25 or greater, and the efficacy of Sertraline is maintained in the patients who have the higher baseline scores.

DR. DAVIS: Could we see those data?

DR. RYDER: I do not have a prepared slide but we may have a photocopy of a SAS output.

DR. CASEY: Perhaps we should save some of these questions and comments for the overall Committee discussion. If we need more data, it would then give the sponsor a chance to collect their thoughts a little bit rather than trying to do it here at the last minute. It would give us a chance to discuss them in full context. But I would give the sponsor the opportunity to address this last question now if they wish.

DR. RYDER: This is not a prepared slide. This is raw SAS output so it is not especially audience friendly.

(Slide)

As I warned you, it is raw SAS output. It is merely a photograph of SAS output.

DR. CASEY: Could you walk us through this?

DR. RYDER: Sure. These are the values for the Sertraline, amitriptyline and placebo groups. They are categorized according to baseline HAM-D severity -- moderate

and severe. The definition for severe is a HAM-D score at baseline of 25 or greater; for moderate, from 17-24; and for mild less than that. There were very few patients in the mild category, as you can gather based on the protocol criteria.

These are the values for the baseline and the change from baseline. For Sertraline, here is the change from baseline in the moderate and the change from baseline in the severe groups, and for amitriptyline and for placebo.

If my memory serves me correctly, this is the intention-to-treat group, last observation carried forward to the final visit.

DR. CASEY: Dr. Davis, is that a response to your question?

DR. DAVIS: I cannot read the slide from this distance.

DR. LEBER: Walk up to it.

DR. CASEY: I believe for the Sertraline moderate group the change was 10.5. In the severe group the change was 14.0. For the amitriptyline moderate group, the change was 11.4 and for the amitriptyline severe group the change was 15.0. The change for the placebo moderate group was 7.9 and for the severe group the change was 8.0.

DR. LINNOILA: Mr. Chairman, interestingly, if one looks at the data, by far the clearest efficacy for Sertra-

line, even though there is efficacy in every group, is in the mild depressive group. If one looks there at the change scores, either as a percentage from baseline or against placebo, clearly, the biggest difference is in the mild group.

DR. RYDER: I do not know if Dr. Linnoila can see it, but the Ns are very small in that group.

DR. LINNOILA: Sorry, I agree.

DR. LEBER: Can I make a suggestion? I would like to stick my two cents in -- I have been quiet until now.

It seems to me that we are engaging in a discussion of a major conceptual area, which is called the size of the treatment effect. I do not believe that there is a standardized method that everyone agrees upon at this point in time, although there are measures that attempt to estimate the size of the treatment effect that everybody agrees upon for a variety of reasons, part of it is statistical; part of it is conceptual; part of it is even what a HAM-D score means. How you map it on the spectrum of disease pathology may vary from center to center, individual to individual and trainee to trainee. So I am not too sure what we are talking about when we look at these numbers in an absolute way.

I was hoping, because I know that Tom Laughren is going to discuss some of these issues, that we could give him a chance to sum up before we get into the detailed nitty-gritty of this. Then, perhaps with that as the framing for

the discussion, we can proceed.

DR. CASEY: That is very good. I agree. Dr. Tamminga?

DR. TAMMINGA: I just have one question about that slide. I want to clarify that that contains no data from the inpatient studies. Is that right?

DR. RYDER: No, this is a subset of protocol 104.

DR. TAMMINGA: Right. Sometime in the course of our discussion, I would like to hear from the Company and from the FDA about the inpatient studies.

DR. LEBER: I believe we will also talk about what the difference is between a positive result, a negative result, a failed study and a null study. All of these are ways of carving nature up to make things understandable in a regulatory sense. We will discuss that, I hope, in the course of the discussion.

DR. LIEBERMAN: Mr. Chairman, if this are the only data which bear on the issue of the relationship of the efficacy of Sertraline to severity of depression, would it be possible to have a copy of that table distributed so that we can see it up close?

DR. CASEY: I do not know if it is technologically possible for the sponsor to produce it. There are a few head nods in the yes direction so I think that we may be able to get a photocopy and have it for our discussions.

DR. DAVIS: Additionally, it would be useful to have data from the other studies that address the same question.

DR. CASEY: With that lengthy introduction, I think we are well prepared for Dr. Laughren's summary. Thank you, Dr. Ryder, for your presentation.

PRESENTATION BY THOMAS P. LAUGHREN, M.D.

DR. LAUGHREN: Thank you, Dan. I would also like to thank my colleagues at FDA and Dr. Ryder for very clear, focused and informative presentations.

(Slide)

I included this slide mostly as a way of reminding you that in addition to clinical data, there are a number of other areas that need to be addressed in an NDA. We have looked very carefully at the chemistry and the pharmacology data. The reviews, at this point, are not finalized, however, based on what we have seen, we see no problems in either of these areas that would preclude the approval of this product.

Of course, it is also required that a drug be characterized from a pharmacokinetic standpoint. Dr. Ryder has presented you with some of the biopharm. data. We have also looked very closely at all of those data and, again, we see no serious problems. There are a few biopharm. issues that still remain to be resolved. So I am going to come back

to that in a bit.

Finally, you have heard a great deal about the clinical data. I want to say a bit more about that, mostly as a way of contrasting what we have gotten with what we would like to get in an ideal setting.

(Slide)

These are the typical bio. studies that we get in an NDA. As I have mentioned, we have, for the most part, received these studies for Sertraline. The ones that we have not yet gotten, we have been promised. These are studies that are either nearing completion or are completed but not yet fully reported.

However, as I mentioned, there are a few areas where we need a little bit more work. We still need to have dissolution studies. We still require additional food effect studies. We are going to want an additional study comparing the pharmacokinetics in morning versus evening dosing. Otherwise, from a bio. standpoint, things look pretty good.

(Slide)

This is just a quick summary of the pharmacokinetic characteristics of Sertraline. You have heard some of these data before. I want to summarize them here because some of these data are relevant to comments I want to make a little bit later on.

As you have heard, Sertraline has a large first-

pass effect. The half-life of the parent of Sertraline is roughly a day. It has been demonstrated to be linear with respect to dose, in the dose range of 50-200 mg which is the dose range which is being proposed for clinical use. As Dr. Ryder mentioned, that has been shown not only with single dose data but also with multiple dose data.

As he also mentioned, there is a very prominent food effect. The availability of Sertraline is increased roughly 40 percent when it is given with food compared to a fasting state, and I want to come back to that issue. It is highly protein bound.

From the data that we have looked at, we have the impression that the clearance of Sertraline is diminished somewhat in the elderly compared to younger individuals. That may also have implications for dosing.

Finally, as has been mentioned, there is a major metabolite, desmethylsertraline, which has a half-life of roughly 60-100 hours. However, it is of importance that the metabolite appears to be relatively inactive compared to the parent.

(Slide)

Now I want to turn to clinical data. I would like to start off by making a few general comments about what, under ideal circumstances, we would like to learn from the clinical phases of drug development. Then I want to relate

that to what we have for Sertraline.

Of course, the primary goals are to establish safety and effectiveness. There is a third goal, which is sometimes neglected somewhat, that is, to develop the data base needed to direct clinicians in how best to use a product. In order to know how to use a product in clinical practice, first of all, one has to know how to dose the typical patient. That means knowing the dose range, knowing, in particular, the lowest effective dose; knowing on what schedule to give the drug; knowing whether or not to give it with food. Finally, if the drug is going to be titrated to some target dose, one needs to know what dose to start with; how to get up there in terms of dose increments; and at what intervals to make those changes. Those are all important clinical matters.

In addition, one would like to know how to individualize the dose for various subgroups, for example, the elderly patients with concomitant disease or patients taking other medications.

A drug which is going to be used in a condition which is chronic, such as depression, one needs to have data on whether or not to continue the drug after one has obtained a response and one would like to know whether or not the drug protects patients from relapse in chronic use. One would also like to know if there are any problems in discontinuing

patients from the drug.

Finally, I think that any drug development program should at least explore the pros and cons of plasma level monitoring in relation to clinical use, both safety and efficacy.

(Slide)

What do we actually have for Sertraline? From the standpoint of efficacy, we have study 104, a 3-way study comparing amitriptyline, Sertraline and placebo, which both we and the firm agree is a significantly positive study.

We also have study 103, a fixed-dose study, which we like in concept. We very much like the dose comparison design. It potentially yields very important information. However, as was noted, there was a problem in the way this study was conducted. Patients were not titrated to dose, leading to high dropouts in the higher dose groups. Nevertheless, we feel that that study is a positive study. It is less strong than 104. As Dr. Lee mentioned, for HAM-D total if the data are reanalyzed excluding certain items that potentially confound the outcome, namely, those items that represent side effects of Sertraline, the overall outcome is a bit more positive.

From a standpoint of safety, the data base of exposed patients is quite large, roughly 3000 patients. Overall, it is quite reassuring in terms of safety. Of

course, that does not rule out the occurrence of rarer side effects that may not show up until a drug reaches a much larger population. But from the standpoint of what we have seen, it looks pretty good.

(Slide)

What about the other issues of establishing a data base needed to instruct clinicians in how to use the drug? Do we know how to dose the typical patient?

This slide is taken from the proposed instructions for use of Sertraline in the proposed labeling initiating treatment. Basically, what this suggests is that patients be dosed in a range of 50-200 mg; that they be started at a dose of 50 mg with upward titration based on clinical response. It suggests that the dose be administered once a day, either with the evening meal or with breakfast, and that patients with renal or hepatic compromise be given lower or less frequent dosing.

I believe these instructions are essentially based on study 104 since that was the study in which patients were titrated in a range of 50-200 mg. There are some questions here. We still do not know what the minimum effective dose is. These instructions do not tell us what increment to use in get patients up to an effective dose. It says nothing about the dose interval, the interval for making the change. From the half-life of Sertraline, roughly a day, we know that

it is not going to reach steady state for roughly a week. So one would want to wait at least a week; one might want to wait a longer. We just do not have the data that bear on that issue.

In addition, it suggests here that it be given with food. One point I would like to make is that we have very little data from the NDA on whether the drug was given with food or not. The discovery of the fairly prominent food effect was not made until well into the development program. So, for example, in studies 103 and 104 we do not know whether patients were dosed with meals.

I do not want to be too critical. I think a very strong mitigating factor for these relative deficiencies is that overall this drug has been shown to be relatively safe when dosed in the interval of 50-200 mg. Nevertheless, these are clinical issues that one would like to have information on.

(Slide)

What about individualizing dose? Again, I mentioned that we have data in the NDA suggesting that clearance may be somewhat diminished in the elderly relative to younger individuals. So I think some change would need to be made in dosing the elderly. If this drug is approved, we will modify the labeling to reflect a need for modified dosing in the elderly.

As far as patients with renal or hepatic dysfunction, as Dr. Ryder mentioned, those data have not yet been submitted so we have not had an opportunity to look at those issues.

I want to mention one other point regarding individualizing treatment. We have no data for this drug in children. Of course, the sponsor is not seeking an indication in children. Nevertheless, I think it is important to point out that depression is an entity that exists in children and if this drug were to be approved, it is likely that some clinicians will want to use this drug in children. I do not want to make a big point of this, except to mention that this is an area where we have no knowledge.

What about the issue of continuation maintenance treatment? Study 320 was intended to look at maintenance efficacy. However, the way that study was designed and conducted makes it very difficult to draw any conclusions about its effect in preventing relapse. It may tell us something about the advantage of continuing patients beyond an initial response but it really does not address the issue of maintenance efficacy.

Despite that, I want to comment the sponsor for doing the study. We rarely have any direct data on continuation or maintenance efficacy in NDAs for depression, even though this is obviously an important clinical issue. So I

think they ought to be commended for doing it. I think the design could have been improved.

As far as discontinuation, we have no evidence from what we have seen that there is any withdrawal syndrome from Sertraline.

(Slide)

What about the final area of therapeutic drug monitoring? These are the only data from the NDA that bear on the issue of relating plasma levels to clinical response. These are data that you have not yet seen. This was an 8-week parallel group dose comparison trial. The dose groups in this trial were 50, 100, 200 and 400 mg. Patients were titrated up to those doses.

We have only very preliminary data from the study, which I understand is still ongoing.

PFIZER REPRESENTATIVE: No.

DR. LAUGHREN: But let me just tell you what has been done here. The sponsor pooled patients and then rank ordered their mean plasma levels. Patients had their plasma levels monitored at various points throughout the trial. So the plasma levels were rank ordered and quintiles were formed, five groups, based on these mean plasma levels. They then looked at the mean change in the HAM-D depression item for each of those quintiles.

What you see is the least change in the low plasma

level group, with monotonically increasing changes as you get to the higher end and with some levelling off as you get the higher plasma level groups. These are very crude data, difficult to draw any conclusions from. But I think it at least suggests that there may be some plasma level effect relationship here that might be worth pursuing.

In conclusion, I think overall the sponsor has provided evidence to suggest that this drug has antidepressant activity; that it can be used in a reasonably safe manner in treating depression. There are some clinical questions that remain -- yes?

DR. LINNOILA: Is this fixed-dose data?

DR. LAUGHREN: It is fixed but patients were titrated.

DR. DAVIS: Were they preassigned to a fixed dose?

DR. LAUGHREN: Yes, it was a dose comparison trial where patients were assigned to a dose group and then titrated to that dose. Again, the dose groups were 50, 100, 200 and 400.

DR. LINNOILA: Thank you.

DR. LAUGHREN: As I have noted, there are some remaining clinical questions that we would like to have answered. We have discussed this with the sponsor. They have committed to doing an additional trial to try and address some of these clinical issues. The proposal would be

to do this on a postmarketing basis. Given the relative lack of important safety findings for this drug, I am inclined to agree with that.

I think I will stop there and let the Committee go on with their deliberations.

DR. CASEY: Thank you, Tom. Are there other questions by the Committee members to be addressed directly to Dr. Laughren? Dr. Hezel?

DR. HEZEL: Is study 86 inpatient or outpatient?

DR. RYDER: Outpatient.

DR. CASEY: Tom, I would like to ask one question regarding the last category on your slide about establishing guidelines for optimal use by clinicians, that is, to inquire as to whether you think that your list of five items should be state-of-the-art, or is it a goal sometime in the future? And do you think, from what we know about the other products that are currently available, could we adequately answer these questions?

DR. LAUGHREN: It is a fair question. Basically, it is more state-of-the-art than what we have for most drugs. It is an ideal goal. It is what we would like to have ideally because it is always a problem, once you get around to deciding that a drug works and that it is reasonably safe, to write labeling. We want to give clinicians as much information as we can to help them in using a drug in clinical

practice. It is these kinds of data that would be very helpful in getting down to the business of writing labeling. I agree that these are ideal goals. For many drugs, if not most drugs, we do not have all of this kind of information.

DR. CASEY: From some of your complimentary comments about the sponsor, I take it that they have done as well or better than many in the past; that it was not meant to be a substantial criticism of the sponsor but, yet, it was not a standing ovation that all is done and we are ready to go on.

DR. LAUGHREN: It was really more a reminder for all of us that there is a great deal that one would like to have before one puts a drug out on the market and gives it to clinicians to use. The sponsor has been very helpful. As I say, I think these are relative deficiencies that are mitigated quite a lot by the fact that the drug is reasonably safe in the dose range in which it has been proposed for use.

DR. CASEY: Thank you. Dr. Leber?

DR. LEBER: I think it should be understood that all comparisons are probably odious. We do not have a comparative efficacy/safety drug law, although, clearly, clinicians using drugs are interested in determining relative efficacy, relative safety and relative utility -- whatever that means.

I think you have to understand that when we face an

application from a regulatory perspective, we are asked to face what the law requires us to do. We are obliged to approve an NDA unless our review finds that the drug is unsafe for use; that inadequate testing has been done to show that the drug is safe. We are required to approve the drug unless we find that the tests submitted failed to contain substantial evidence of efficacy. That means more than one investigation which is adequate and well controlled which would allow experts -- experts by experience, training and background -- to reach a conclusion that the drug is effective. And we are obliged to approve the drug unless we find that the labeling is false or misleading in some particular.

Now, those are three of the seven items in the law. But, given that perspective, you can understand why we have to look at the application submitted to us and recognize, in a way, that we can exhort people to do more. But the law did not set out a very Draconian or Procrustean set of standards that have to be met. So a lot of what Tom has said is something we seek. That is sought in the sense that "A man's reach should exceed his grasp, or what's a heaven for?" rather than necessarily setting it out as a demand in a regulatory way because each time, with your help, we have to face what the firm has submitted in a concrete way and decide whether it meets the test of law.

I can guarantee, things being what they are in the

world, that the ideal is never attained. As drug development programs go, you have to ask yourselves whether this firm has provided you with enough information to reach a conclusion for the law's purposes. If you were academicians talking about the idealized world, I am sure that each and every one of us could spend many hours talking about designs or things we would like to attain.

But I think that is an important point of the regulatory charge to this Committee -- understand the question we are asking and the kind of advice we need. I know that many things have not been done with this drug. Some of them may arise and the reasons they have not been done. For example, that patients who are in-hospital show a response has not been documented. But you have to ask the question what does it mean not to document that? What does it mean today to say that you have hospitalized depressed patients? And what inference can you draw from small studies which fail to show that? In a regulatory sense, it is distinct from other general senses.

So if we can come back to the regulatory flavor of the questions, I think it would be useful.

DR. CASEY: Dr. Escobar, do you have a question for Dr. Laughren?

DR. ESCOBAR: I want to ask about previous history. Has it happened that an antidepressant is approved without

evidence of inpatient efficacy?

DR. LEBER: Yes.

DR. ESCOBAR: It has?

DR. LEBER: Many times. Again, you have to make very certain what you mean when you say evidence of inpatient efficacy. I am going to ask you a question. Let's assume that we did a study with inpatients and we found in that study that we had a difference that achieved statistical significance between those randomized to the experimental treatment and those randomized to placebo. Would you reach a conclusion that the drug is an ineffective treatment for inpatients on the basis of that?

The head is being shaken in a negative direction, for those of you who cannot read that on the transcript.

I think that the whole issue that I thought we might get into and that was being raised earlier is the issue of the size of the treatment effect and the meaning of that treatment effect as extrapolated to the world of use. I have no idea what constitutes proof of efficacy, except on the basis of what we, as a Committee, agree on an as ad hoc case as there needs to be. You can be guided by the past but the inference is an abstraction -- what is an antidepressant?

I think over the past 27 years or so since people have been looking at that question, we have taken changes on the HAM-D, the Clinical Global Impression of severity, POMS

factors and a variety of other things and taken those as testimony or indicators of efficacy. But that is tradition. That is not truth. Anyway, that is the answer to that question.

DR. CASEY: I will take those comments as the opportunity for transition to the next section of our meeting, which is to turn it over for discussion to the Committee members. I will point out at the beginning of our discussion that we have two questions put forth before us by the Division:

One, has the sponsor provided evidence for more than one adequate and well-controlled clinical investigation that supports the conclusion that Sertraline is effective for the treatment of depression?

Two, has the sponsor provided evidence that Sertraline is safe when used in the treatment of depression?

Those are the two general issues before us. Now I turn the meeting over to us as a Committee for discussion. Let's start with the first issue, have they shown efficacy in more than one study that is adequate?

DR. HEZEL: My question is sort of related to the inpatient/outpatient issue. There were no plasma levels drawn on the outpatient studies. Is that correct?

DR. LEBER: We will have to ask the firm. Do 103 and 104 have any plasma levels?

DR. RYDER: No plasma levels were determined in protocols 103 and 104.

DR. CASEY: The answer from the sponsor was that there were no plasma levels from outpatient studies.

DR. RYDER: The single slide that Dr. Laughren showed was from study 86. That was an outpatient trial, not an inpatient study.

DR. HEZEL: But 103 and 104 that we have all the data on had no plasma levels done?

DR. RYDER: Plasma levels were not drawn in those studies.

DR. HEZEL: And those were outpatients.

DR. RYDER: That is correct.

DR. HEZEL: So we relied on patient reports as to whether or not they took the drug?

DR. RYDER: Pill counts were done as a measure of compliance.

DR. HEZEL: So each time they visited the clinic, they brought in their pill container for pill counts?

DR. RYDER: Correct.

DR. HEZEL: Okay. The problem I have with that is that on the inpatient studies where drugs and environment are controlled, we did not see a significant effect. But in the outpatient studies where drugs and all other variables are not controlled, we saw the therapeutic effect. The difficulty

I have with that then is the leap of faith that I must attribute the therapeutic effect to Sertraline.

The literature is full of information about how poorly we all take prescribed medications, whether or now we take them on time, at all, finish the prescription -- that is quite common.

DR. LEBER: Does anyone want to answer that question? I think I can start an answer but it is not an easy one. Let's start by talking about if patients in any of the controlled, randomized trials that do show evidence of efficacy had failed to take their medication, do you think that would increase or decrease the size of the treatment effect? I am asking you, Dr. Hezel. If this is going to be a dialogue, I think we ought to get into it.

DR. HEZEL: On the outpatient studies?

DR. LEBER: On the studies that have shown -- remember, we are treating these as fixed effects. That means that we are not making any generalizations about this. We are saying, in terms of a randomized, controlled trial that contrasts between 4 groups in 1 study and 3 groups in the other study, and we find a difference between treatments in a randomized, controlled trial and we attribute it to perhaps 3 or 4 different sources: It could be fraud, which we would not think likely because these are multicenter trials in [illegible] It could be

chance but we do statistics to rule out the likelihood that that is claimable. It could be bias, some systematic bias in the design and I suppose you are arguing that I could take treatments in such a way that I would be seeing a placebo effect only and that is what is giving me the results. Or it could be a treatment effect due to the drug.

I think if you think of those 4 possible explanations for the randomized, controlled trials that are positive, the most reasonable explanation for them -- dealing with those studies per se -- is that the drug did it. In the 3-way study with amitriptyline there is a consistency that both active treatments showed a difference from placebo that is favorable. In the study comparing 50, 100 and 200 versus placebo there is again a consistency, although not every level reached statistical significance but across every center that you look at in those studies there is basically the same directional trend. If you took that as a whole, you would believe, based upon what accounts for the differences, that the drug must have done it.

The issue that I think most people are dealing with is, is the size of that effect significant enough in terms of clinical meaning to allow us to conclude that we have a bona fide antidepressant that has clinical utility to go into the armamentarium? But I think in terms of the internal validity of those studies, that is, their differences and what they

could be attributed to, most reasonable people would agree that they would have to be attributed to the effect of drug.

Again, I would welcome hearing from the Committee if anyone would choose to interpret that it is not due to drug, and hear the reasons for it.

DR. ESCOBAR: I guess the reason I worry about outpatient versus inpatient data is because there is some evidence that a sizeable number of patients recruited in the outpatient clinic for this type of trial is selected through newspaper advertisements. They are so-called symptomatic volunteers. Patients who are recruited in inpatient units are what we may view as "psychiatric" patients.

In all fairness to Sertraline, this goes beyond the individual agent that we are talking about today, but what often happens is that once one of these agents is approved on the basis of outpatient data, then when you begin using it in the traditional psychiatric populations you may find some surprises. So my concern about consistency in outpatient and inpatient data is because I have a feeling that these populations are very different.

DR. DAVIS: I wonder if it would be possible to get printouts for the other studies, like 103 and 315, where there is a breakdown for severe versus moderate depression. That would speak to this point about symptomatic volunteers. If psychiatrists are going to use the drug, they need to know

if it will help the patients with moderate to severe endogenous depression, which may be a different animal from mild depression. An approximation for that would be the breakdown of patients by severe and moderate.

DR. LEBER: Can I ask something of the firm? I want to make very clear that in 103 and 104 there are not symptomatic volunteers, as I understand it. That is a question to the firm.

DR. RYDER: There are not symptomatic volunteers. These are patients with DSM-III diagnoses of major depression. If you would like, Dr. Mendels is here and he was one of the protocol 104 investigators --

DR. ESCOBAR: The questions is whether they advertised for the subjects.

DR. RYDER: Some of the clinics, I imagine, advertised. I do not remember systematically collecting those data.

DR. LEBER: Nonetheless, there would be a difference between the symptomatic volunteer and a patient who is depressed who finds out that he can get free treatment for depression. It is not unique to firms. Doesn't the NIMH do it?

(Laughter)

DR. CASEY: Dr. Mendels, would you like to respond?

DR. MENDELS: I can make a couple of comments in

terms of the several issues that have been raised as part of this discussion. First of all, insofar as the inpatient studies that were completed and were a part of the NDA, I think it would be fair to say, as an outsider to the firm, that these were not very good studies. There were small numbers. They were done at a very early stage. When you look at the sites at which they were done, they were less than optimal sites. So it would be my conclusion that these sites may not have had good patients with major depressive disorder. They would be more likely to have chronic resistant type patients.

So I, personally, do not think that those are truly negative studies. I do not think those two studies suggest that the drug does not work. I think those were just studies in which probably very few drugs might have worked.

Secondly, I think the term symptomatic volunteer obviously has a stormy history in our field. Without wanting to get into that whole debate now, I think it is fair to say that we know from NIMH and other studies that somewhere on the order of 70 percent of people with major depressive disorder in the community do not present for treatment in traditional psychiatric or medical settings. A significant percentage of these 70 percent are being captured at a number of clinical trial sites. There is no doubt about it. Clinical trial

I think it is pejorative to suggest, however, that these people do not necessarily have major depressive disorder. I think all of the investigators who were involved in these two studies are reasonably experienced psychiatrists who have worked in a number of settings, many of them in major academic institutions. I think it would be fair to say that the diagnosis of major depressive disorder was based in DSM-III or was made for these patients.

The ways in which doctors attract patients today are changing very rapidly. We see increasing numbers of physicians in traditional practices, including many hospitals, who are advertising for patients. I think we have to begin to review that issue. The critical question is not so much where the patients come from but are the patients alerted to the possibility that they have a condition for which a treatment or a potential treatment is being offered?

Finally, and I am hoping you will be able to see the data, there has been a reanalysis on several of the studies in which the patients were broken down according to HAM-D scores above 24 and below 24. It is my understanding, having seen the data and, obviously you will want to see them for yourselves, that there is a fairly consistent pattern across studies which suggests that the patients with the highest scores do at least as well as the patients with the lowest scores and, in some instances, perhaps even a little better.

Thank you.

DR. CASEY: Is it possible that the Committee might see that data reanalysis sometime today?

(Dr. Ryder nods in agreement)

The answer is yes. So I imagine a photocopying machine somewhere is warming up. Dr. Linnoila?

DR. LINNOILA: I think that the question about the inpatient versus outpatient is an important one because, first of all, it is important for clinical practice. We know that mental health insurance coverage is not very great and it is not good to put patients on inefficient treatment. If there is a real indication that in some severe inpatients certain drugs work less well, then I think that that is important to know for medical practice.

The second issue is that there are data from Europe, from large, multicenter studies looking at several serotonin reuptake inhibitors versus conventional tricyclics, which fairly convincingly suggest that the tricyclics have somewhat higher response rates in the patients than the serotonin reuptake inhibitors. However, that comes at a price. There are clearly more side effects to the tricyclics than there are to the serotonin reuptake inhibitors. Typically, it is not a random selection as to who ends up being an inpatient versus who ends up being an outpatient. The inpatients tend to be the most severely ill, more

difficult to manage on an outpatient basis, whether they are suicidal, whether it is because they do not sleep, they have significant cognitive impairment as a result of their depression. And I think that this, again, is an important clinical issue for which it is important for physicians to know about the efficacy of drugs or lack of efficacy of drugs.

Finally, there is no question that serotonin reuptake inhibitors are very popular antidepressants currently throughout the Western world and they are clearly effective. But one has a seat of the pants feeling that they are particularly effective in patients who perhaps fulfill criteria for dysthymia more than major endogenous depression. And that is the question which one would like to get more information on.

DR. LEBER: I have one concern about the word inpatient. You describe very clearly, Markku, a set of criteria that you might use to hospitalize a patient under a set of circumstances where you are in control of hospitalization. I think those of us who may have worked at different hospitals recognize that in some places inpatient hospitalization is a function of resources, private payment and a variety of other things.

There is a certain heterogeneity to the nominalism associated with it. It is conceivable, as I was making the

point when I was asking the question earlier, that the hospitalized state may be a signal or a marker. But I do not know what percentage of the variance of severity, or difficult response, or difficult to manage is actually captured by that.

The second thing I think you have to worry about in this particular case is sample size variance and the size of the treatment effect which, as we all know, determines the statistical significance. If you do power studies and power analyses before you conduct studies, obviously you need a lot of patients when you look at treatments of small or modest treatment effect sizes. Generally speaking, in order to collect sufficient numbers of patients to detect drugs with treatment effect sizes of this magnitude, you are pretty much stuck with outpatient, ambulatory studies. All we are acknowledging in doing that is saying that the size of the average treatment effect, which I have yet to define, tends to be such that if you want to document the efficacy of most antidepressants in a robust way, you are going to have to go outpatient in a multicenter trial.

Now, we have that data set from large multicenter trials and we know that the drug is effective in that setting. Now you go inpatient and you do small studies -- I do not know what the variance looks like but the N is small. Can you hope to achieve statistical significance even if you

got the same effect treatment sites? I doubt it. They are grossly under-powered.

Under those circumstances, to reach a conclusion that the drug does not work inpatient is not really fair because the question is what kind of a test. You may say you would rather use ECT because ECT is going to give you your biggest bang over the shortest period of time. But, as I said earlier, we are not involved in a comparative question necessarily. It may be useful for directions for you. It is interesting to everyone in learning how to use the drug. But it is not a sine qua non for approval that you work in inpatients. That is my explanation.

DR. CASEY: To address this issue a little bit more, I have asked Dr. Lee if she could make a few comments in a few minutes, when she has a chance to collect her thoughts and her figures, about the power of those inpatient studies and the ability of those studies to find an effect if there were one, and to look at some of the weaknesses that may have been involved with the methodology.

DR. LIEBERMAN: I think --

DR. CASEY: Excuse me, Jeffrey, is your comment about the power or the inadequacies of the inpatient study that Dr. Lee is going to include in her address?

DR. LIEBERMAN: It has to do with the therapeutic effect and severity of depression relationship, which is, in

part, related to by inpatients versus outpatients. Do you want me to talk about that or to refrain until after she has made her comments?

DR. CASEY: You be the judge.

DR. LIEBERMAN: Well, let me speak quickly, which is to say that in terms of the first question that we are charged to consider, I think the evidence, limited as it may be, does speak to the efficacy of Sertraline as an antidepressant in an outpatient population. The question that has been raised and the question that lingers in my mind is whether Sertraline is effective in more severe forms of depression and what the dose-response relationship is in that population relative to what the dose-response relationship is to adverse effects in those patients.

From the data included in the studies that are accepted as methodologically adequate, it looks as if Sertraline has a moderate therapeutic effect which is superior to placebo and comparable to that of amitriptyline, although numerically it is less robust than with amitriptyline. It is also apparent from the data that there is a relationship between side effects and dosage of Sertraline, to the extent that the dose range is reflected in the patient sample.

It is implied that in more severely ill patients higher doses of the drug may be warranted. Thus, the

question is what is the relation between therapeutic dose response and adverse effect dose response at the higher dose ranges?

In the two inpatient studies which did not demonstrate efficacy, the dose range went up to 400 mg. Does that imply, by virtue of lack of efficacy, that this was no better? The data were not detailed enough to enable that interpretation. But that would be one aspect of this relationship between therapeutic effect and severity of illness that I would put forward, namely, what is the relationship of the dose-response curves for therapeutic and adverse effects in the more severely symptomatic group?

DR. CASEY: Dr. Leber has one comment.

DR. LEBER: Again, this is a methodological issue that I would like to raise once again. It goes back many years to the concept of assay sensitivity, which Modell and Hood introduced and widely circulated in 1958, in an article they wrote on behalf of the Council of Drugs for AMA.

This was based on pain studies and they pointed out that there was no point ever discussing the results of a clinical a trial unless you know that that clinical trial has the ability to discriminate an active drug from placebo and can, in fact, discriminate between two active drugs if you are making a comparative analysis.

So if we have studies such as those that failed to

find a difference between active drugs or those that failed to find a difference between an active drug and placebo, we do not have assay sensitivity in that trial. Generally, as an Agency, except in areas where we know historically that the results could not have happened in the absence of drug treatment, we have tended to take such studies and treat them as uninformative. They do not tell you much.

If you had an inpatient study where you randomized three different treatments, including the standard used in the world, whatever that might be, IMI or AMI, the new drug and placebo, and you showed time and time again that the standard drug beat placebo with big treatment effects and your drug did not, I think you would have what we would describe as a study with assay sensitivity which failed to show the comparative efficacy of the new drug.

But if all you have are studies that fail to show differences, I think that the interpretation from the point of view of a regulator would be that it is sort of like using a spectrophotometer that is not recording anything or is recording ODs of 3. You cannot interpret that. That does not mean anything.

So my question for the firm, before we trounce them for not doing inpatient studies, is have you got any three-way inpatient studies that compare a standard drug with placebo? If so, how did the standard drug do? So that would

be my challenge before you get into power considerations.

DR. CASEY: That is a question to the sponsor.

DR. RYDER: The only inpatient studies were the two that were presented. They included Sertraline and placebo. There are no inpatient studies with an active agent.

DR. LINNOILA: It is always very nice to hear the spectrophotometer analogy but why is it that the operator of the spectrophotometer is always left out which, in this case, is the designer of the studies?

DR. LEBER: I do not understand what you mean, Markku. You had better explain your metaphor. Why don't you explain it because, obviously, you have something --

DR. LINNOILA: Well, I have some problems because we say they do not for one reason or another and there are certain deficiencies. I think it is deplorable that there are certain deficiencies because that basically means that you are asking us to evaluate data which do not stand up to scrutiny and --

DR. LEBER: Who is asking you to evaluate the data? The firm submits an NDA. They are responsible for the drug development plan. Are you saying that what they did was deplorable?

DR. LINNOILA: Well, I am saying that some of the inpatient studies clearly are not well designed, for one reason or another, if they do not yield the answer.

DR. CASEY: They yield an answer. The question is what is the meaning of the answer. Dr. Lee will help us interpret some of that. Dr. Lee, I am sorry for the long delay.

DR. LEE: No problem. Both inpatient studies followed the same design. They were both fixed-dose studies without titration, just like protocol 103. There were 5 treatment groups in each trial, the Sertraline 50, 100, 200, 400 mg and placebo. The duration of the trial was only 4 weeks. They needed a score of 18 to get in. One protocol was carried out in the U.S. There were 11 investigators who enrolled 132 patients. So that ended up with a total of 25 patients in each treatment group, which meant that each investigator might have had 2 rounds, 10 patients that he entered into the study, which is not very much. That is a very small number.

In protocol 310, which is the same design, carried out in the U.K., there were 16 investigators and 174 patients. That was about 35 patients per treatment group. Still each investigator only had about 2 patients per treatment group.

In the U.K. study, much like protocol 320 which was also carried out in Europe, there is a large number of concomitant psychotropic medications, which makes it even more difficult to show a difference between the drug and placebo. Those are all the data I have with me.

DR. LEBER: Hillary, did you look at the HAM-D depression item in those studies as well as the total?

DR. LEE: When they were not significant, I just went over the design.

DR. LEBER: I want to bring up one point that I thought Tom was going to cover. It has to do with the outcome assessment using the HAM-D total rating scale for depression and whether, in fact, the HAM-D registers the side effects of a drug. Hillary talked about it. I think it is something that people have complained about before.

When Max Hamilton designed this scale in 1960 or before that, he thought it would be used by experts in the treatment of depression. He never thought he would be designing a scale with selectivity or specificity. He assumed that the scale would register the phenomena that were common in endogenomorphic depression, or whatever they called depression in those days that was severe. He did not design it as a differential test. In fact, if you read his paper he said that he expected that the highest scores were in people who truly had depression but warned about using it in populations that did not have depression because it would register other phenomena. Obviously, he is talking about the vegetative signs, anxiety, insomnia, GI disturbances and the like.

We have come forward almost 30 years and we begin

gg

to have drugs which produce side effects which register on the HAM-D total score. I think Hillary has done some analyses and she showed them in the aggregate presentations. The HAM-D total score did worse. So the real test of this might be that if you had an inpatient drug and you are watching as a clinician, what are you looking at? You are looking at vegetative signs, agitation, insomnia. You hit them with a drug which makes them agitated, upset and so on. You may not give the drug an adequate chance, especially in a fixed-dose design.

I think these are terrible studies the way they worked out. I agree with Markku. But I think before you say a negative study means that the drug does not work, you have to examine whether it is a fair study and that is different, reaching a conclusion that it is a bad study and did not test the issue, from concluding that a drug does not work in inpatients. We do not know that yet.

DR. CASEY: To summarize Dr. Lee, I think, in fairness, she said there were too few patients with too many investigators and too much variance to give us a meaningful answer about the efficacy of the compound in the inpatient studies.

DR. LEE: There is also the problem of design. It was the wrong design.

DR. LINNOILA: Mr. Chairman, if I heard the numbers

correctly, there were 300 patients in the inpatient studies. Did I hear you correctly?

DR. LEE: Yes.

DR. LINNOILA: So then it comes exactly back to the point that the spectrophotometers in different centers were not calibrated, using your analogy. They did not measure the same thing. If you do not have power with 300 patients, then something is wrong.

DR. LEBER: I think this also requires a point of explication. There is a difference between power issues failing to show an effect and issues of design. You could have a thousand patients in an active control study and it would not make any difference because you may not have a population of patients who are sensitive to the drug, or the conditions of use may be wrong. That is the reason why I think Modell and Hood were talking about the need for a measure of the sample's sensitivity to the drug. It is not just the drug; it is the sample.

Without having a standard drug in there, you have no way of candling the value of the studies that were done even if they had enormous power. It is not just statistical power; it is a design that fails to tell you that even an active standard drug would have worked.

In past NDAs that have come before this Committee, we have had multiple three-way trial designs which failed to

distinguish the standard drug from placebo. We have been burned before -- we had nomifensine and one of the six three-way trials showed a difference between standard drug and placebo. In another big drug that has yet to come to this country, marketed elsewhere in the world, a number of studies have failed to discriminate their drug from placebo. Failed studies in depression are common. Why they are -- ah, we all have our reasons and ideas but I think you have to accept that as part of the topography that we are working with.

I sympathize with you. We would prefer to have better studies but we are stuck with the applications we get. We try to shake them a bit but we cannot control them.

DR. HAMMER: Can I toss in a few statistical issues as the statistician here? In a sense, what we have is not just an effect size problem but, in a sense, what we are trying to wrestle with is sort of an intuitive multiple comparison problem. If all we had was the two outpatient studies and they fairly clearly showed some Sertraline effect, we would have what I interpret as the criteria necessary for us to say go ahead and approve the drug. That is, we have more than one well-controlled study that demonstrates an effect.

So the question is how do we interpret those two positive results in the context of several more studies that fail to demonstrate that effect? I am not sure I have an

answer to that but I am not sure that the law requires me to have an answer to that -- fortunately or unfortunately. That would mean, in a sense, that the sponsor could just do studies until the cows come home until he gets two of them that are statistically significant by chance alone, walks them out and says that he has met the criteria.

I have also participated in running and designing a number of depression trials and they are real difficult to do and it is difficult to demonstrate an effect.

Perhaps we might address some of these issues simply with labeling -- you know, they have failed to demonstrate an inpatient effect. That is a different thing statistically from demonstrating that there is no inpatient effect. Perhaps the labeling should state that they failed to demonstrate an inpatient effect.

DR. CASEY: Could I suggest a slightly different answer? While you were talking I was in collaboration with Dr. Laughren and we came to a similar conclusion, that is, maybe the Committee should address the language as something saying that the effect of Sertraline in inpatients is inadequately studied, period. That is our conclusion, isn't it? Is it? I think it is, that it is inadequately studied. It is a one-sentence line that goes in labeling.

DR. ESCOBAR: The question is does it have to be adequately studied?

DR. HAMMER: But given that it has been studied, perhaps the labeling ought to reflect that information.

DR. LEBER: Labeling of current drugs does say that they have failed to demonstrate an effect in a particular area. But to make a moral judgment about inadequacy or adequacy I think would be hard for us to put in labeling.

DR. LIEBERMAN: Just to frame the question again, I do not think anybody is saying that there is a lack of efficacy. It is a question, in my mind, as to whether the efficacy is comparable across all levels of severity of illness and, to some extent, that overlaps with the patient status, inpatient versus outpatient, and whether the efficacy at the doses required in different levels of severity has the same side effect risks involved. For those questions the data are not sufficient to enable us to come to a clear interpretation or conclusion.

DR. LEBER: By the way, I have to correct myself. Tom, do you want to read what we already allowed for Prozac? You ought to get it while it is hot.

DR. LAUGHREN: I have the labeling for Prozac. Incidentally, we had a very similar problem in making a decision about Prozac. We did not have data from an inpatient study that demonstrated that it was effective. As an approach to that deficiency, there is a statement in the indications section for Prozac: The antidepressant action of

Prozac in hospitalized depressed patients has not been adequately studied. I think we are basically in the same situation here.

DR. PRIEN: In terms of the regulatory charge outlined by Paul that the sponsor provided evidence from more than one adequate study that Sertraline is effective for the treatment of depression, this group seems to agree that for outpatients we have established efficacy. I am not so sure. I would just like to throw this out because, as Paul states, there are no set boundaries, no standards for what constitutes a clinically significant difference or treatment outcome. So you are dealing with state-of-the-art opinion.

I think study 104 demonstrates an effective treatment difference between drug and placebo. I am somewhat less convinced about study 103, where I believe that the differences in mean change scores between drug and placebo are marginal perhaps, at best, in looking at all the outcome criteria scales.

I may be persuaded otherwise if I could see that the reanalysis of the HAM-D in terms of items that may reflect side effects might show something different. But I think this is something that the Committee has to deal with, that is, whether they actually think that the efficacy data are sufficient to classify it as more effective than placebo, not only statistically but clinically.

DR. CASEY: Dr. Prien has given us the challenge to review study 103 related to its clinical efficacy.

DR. TAMMINGA: It is a bit unclear to me how to separate statistical efficacy from clinical efficacy.

DR. HAMMER: I can address that. With sufficient subjects, a very small difference between several groups may well be statistically significant but for the size of the effect the difference between groups may be so small as to be unimportant in terms of treatment. Suppose I were working on some sort of a drug that influenced survival in cancer patients and I could demonstrate with sufficient subjects that the drug increased life expectancy from 435 days to 435.2 days. That might be statistically significant but it would be clinically unimportant.

DR. LEBER: You also have to understand, however, from the point of view of regulation and by tradition that we have not attempted to set a minimum size of a treatment effect that must be achieved. For example, the suggestion of using a HAM-D 50 percent change score and then doing a categorical analysis of percent meeting that, most people, I guess, would be willing to say that a 50 percent change from HAM-D is some evidence of effectiveness. The answer has always been why not 45 percent or 65 percent and so on?

So for whatever the reasons are, historically we have chosen to allow people to document the efficacy of a

drug without ever dealing explicitly with the question that Bob has raised. That does not say that it is not a good question. The problem is how to reach consensus on how to judge that.

Remember, the HAM-D is a collection of items. Everybody is rating those items on the basis of an idiosyncratic, self-contained measure. We have seen people with symptomatic volunteers get HAM-Ds of 30 and more. We have seen other people, trained in different environments with different standards and experience, get very low HAM-Ds in very sick patients.

So the first question you have since you are not using an absolute measurement like degrees Calvin is what the measure means. The second thing is, if you do have a measure which is arbitrary and scaled to the world in a strange way, how bit a change is enough?

The only answer I know to this has really been provided by statisticians who tend to look at the size of the average treatment change in terms of the size of the standard deviation of the population that is being tested. It is called the standardized difference. I believe Cohen has used it in power studies, and there is a variation on that known as the "small f" test, in which he looks at the standard deviation of the cell means versus the standard deviation of the population and characterizes treatment standards that

way.

DR. HAMMER: But even those do not address the issue of clinical importance in a sense, which is a substantive question not a statistical question. You can use the statistics to answer, in a sense, the substantive question once you have decided what the appropriate substantive question is.

DR. LEBER: But they do give you a historical estimate of what we are used to seeing in terms of the size of the treatment effect, that is, say, the max mean minus the smallest mean over the standard deviation of the population.

Just for kicks, this morning I ran a few of these using a program I got from a colleague, Dr. Lieberman, based on Cohen and it suggests on the HAM-D item, getting effect sizes for the "small f" of about 0.25-0.2, which he would classify as a modest to minimal treatment effect size. That is from studies 103 and 104. So they are not big treatment effect sizes, make no and's, if's or but's about it, even defined that way. It is a tough problem.

DR. LIEBERMAN: In terms of the clinical efficacy issue, Bob, is your question regarding study 103 about the magnitude of the placebo response rate or simply the quantitative difference in the Hamilton or CGI point reduction between Sertraline and the placebo groups?

DR. PRIEN: Both. But I wanted to add one thing to

what Paul has said. There is one safeguard actually that I think is not often applied with some of the pharmaceutical company studies in trying to estimate whether a statistically significant difference is clinically significant. In choosing sample size, I know that grants that come in for review are very carefully reviewed so that you are not using too large or too small a sample size for the difference that you consider to have an acceptable treatment effect.

Unfortunately, in some pharmaceutical company studies, and other studies in the field, they pour patients in from 15, 20 centers with very large sample sizes. If those samples were determined a priori on the basis of need and need alone to determine the validity of a treatment difference in the population, I would be a lot more secure about the findings that are obtained.

DR. LINNOILA: Furthermore, if we talk in very loose terms of clinical versus statistical significance, I think that in the clinical realm it is important to note that there was a very high dropout rate in the study where Bob raised questions. We can again come to the issue and say that there was a design flaw because they started everybody at the high dose. But I do not think that we are here to make excuses for flaws in designs. We are here to evaluate data and if the data are poorly collected, for whatever reason, they have to speak for themselves.

DR. LEBER: Just to stir things up a little bit more, any NDA application is going to have hundreds of volumes and many thousands of patients and a lot of studies. Several of those studies are going to be botched. I guarantee it. You know, certain planes will crash; certain trains will come in late; certain things do not happen.

We are always in a position of trying to make a fair judgment, knowing that we have to weigh the requirements of the law, the expectations of a public that wants freer access to new and effective drugs, even if they are not necessarily as potent on a milligram basis or even in terms of the size of the treatment effect as others. They may have other advantages. It is very hard to judge a drug on a single dimension.

We take the data base as we have it. I think our staff this morning a very nice job of the kind of candling we do. At the end of that, we have to go back to the regulatory charge that I raised. It is not that they are entitled to every claim, every superlative ever made, but is the application, as submitted, such that we have a right to conclude it does not have evidence of safety for use; it does not have evidence of efficacy or it is inadequately labeled. I am not counting now the chemistry requirements.

If we can reach those conclusions, we can reject the application. If we cannot reach those conclusions, you

have to approve the application. That is where we are. As a Committee, you might say to us, "look, we think the standards in this field are terrible. People have been getting away with non-substantive efficacy for years. We'd like you to change your standards. We think that henceforth you're going to need one standard deviation worth of change on the means in order to get it or some other rule."

Fine. Where do we go from there? What about all the drugs that are out there? How many of them can meet that standard? Can we enforce it legally, and so on and so forth?

So I would be delighted to hear discussion about how we can deal with it. Bob can ask the question about the size of the treatment effect but if you tell me a practical way to answer it, I would love to hear it.

DR. CASEY: Dr. Prien has an answer.

DR. PRIEN: Let me throw this back to you, Paul. What would you do if, say, you get a statistically significant different with a difference in a HAM-D of perhaps one-tenth or three-tenths of a point; a CGI score that is just slightly different? What would you do with this, Paul? Are we going to fall back to the position that statistical significance is enough to demonstrate efficacy or are you going to leave this in the hands of a committee to try to judge whether, in their opinion, this is a clinically effective difference?

DR. LEBER: Again I think the answer is that we do

turn to committees. You set up an extreme case where we have used 2000 patients to demonstrate a statistical difference at 0.2 units. But, actually, if you look across the size of the treatment effects here, with all the flaws, they are not so out of line with what we see in drug sees that show efficacy of previous antidepressants. I did not say there are enormous effects but our problem is never saying the minimum size of effect. It is a gestalt. By the way, that is a way of escaping dealing with this in a very careful way. It is the judgment, the sentiment almost, of individuals about what they see the size of the treatment coming out of the studies versus the risks of a drug versus all of it put together in some private way. You, as experts, offer an opinion about whether you find it convincing.

If you can articulate this into an algorithm that is quantitative and specific, I would love to see it.

DR. LIEBERMAN: Can I get a word in here?

DR. CASEY: You can try. Dr. Lieberman?

DR. LIEBERMAN: Just to respond very specifically to the question that Bob has raised about study 103, it would seem that the treatment effect is diluted to a significant degree, as Dr. Lee pointed out, by the dropout rate in the higher doses. Assuming we were going to see a more robust treatment effect associated with the higher doses, that was lost because of the increased dropout rate due to whatever

factors led to that. As a result, the overall effect is diminished somewhat and is most consistently apparent at the lower dose group which shows efficacy, but scaled down in magnitude, and is probably to some degree responsible for what we are perceiving as statistically significant but is not necessarily robust enough to be clinical effects.

In study 104, which is a single dose versus amitriptyline, there is a more consistent therapeutic effect that is observed.

DR. ESCOBAR: I guess there are some standards in epidemiology. At least I remember that unless you are able to interview 75 percent or more of the target population, the data were going to be very questionable. Is there such a standard for psychopharmacology? Do we have to have 50 percent of the people complete before we take it seriously?

DR. LEBER: There is some history on this. Everyone who knows about it is laughing because at one point we arbitrarily said, "look, don't examine a study after it has lost more than 30 percent in any treatment arm." That was met with derision and attack and all sorts of other things because it had a lot of bad statistical properties for people who wanted to show drugs worked and they had to go until there were 10 percent left. So we have never been able to handle that issue.

What you saw us do here was to take a look at the

observed cases and the endpoint analyses at each point of time, and we tried to understand the data. I think what you are doing, Jeff, is exactly that. You were saying, "look, if you have a lot of early dropouts from the high dose group, 200 mg Sertraline, we are carrying forward to the average score in LOCF analyses people who are very sick who, even if they got placebo, would improve over time had they stayed in. So you are biasing that high dose study against the drug.

If you look at the observed cases, on the other hand, it is usually biased in favor of placebo because the patients who do not do well have been kicked out because they are not doing well and in the study we are allowed to drop people out. So observed cases raises the placebo response rate higher than it ought to be. So all of these things make the differences smaller than we think they would have been had the studies been carried to completion according to the design.

We never have studies without dropouts. Usually we are worried about studies where there is a differential pattern, as there was here. I think the statisticians, Dr. Lee and the clinical presentation nicely showed that in the 200 mg group you got overwhelmingly large numbers of dropouts very early on. It is a flaw because of the method of induction on the drug. But it allows you to understand the data in terms of the final LOCF for that group being heavily

biased by early dropouts and you know that on average in depression people's scores are improving. So I guarantee, by subjective assessment guarantee, that those LOCF analyses for the 200 are lower in the size of the treatment effect illustrated than they might have been. I think that was your point.

DR. PRIEN: I agree. I think, having milked this issue a bit, the differences in mean change scores were marginal. However, what would sway my opinion toward it being an effective drug is the point that you raised, that the percent improvement and the final last observed or observed cases scores were similar to those you attain with other antidepressants. I think what you are seeing is a higher rate of improvement with placebo groups, which seems to be relatively more prevalent now with studies than it used to be, for whatever reason. But seems that the drug efficacy outcomes seem to remain relatively constant. But if you look back five or ten years ago with respect to placebo results, you are seeing a significant increase in the improvement rate of patients on placebo with the measures that are being utilized. What this means, I am not sure.

DR. CASEY: I have the feeling that we are coming to consensus that there are some things we do not know.

(Laughter)

I would like to try to get us to move forward and

see if there is any more discussion on this issue to see whether it would be time to address the first question about efficacy. Dr. Davis?

DR. DAVIS: I would like to ask for comments from both the FDA and the drug Company about the dropouts for lack of efficacy, as on page 69.

DR. CASEY: Page 69 of the sponsor's submission?

DR. DAVIS: Yes. For Sertraline there was 11.4 percent dropouts for lack of efficacy; placebo, 22.4; active control, 5.9.

DR. CASEY: The specific line in question is lack of efficacy. That is approximately half way down the table.

DR. DAVIS: Yes.

DR. LAUGHREN: This is a pooled data set across all multiple dose studies. Notice that the sample size for Sertraline is 2710. I believe some of these were even open trials. It is every patient in the development program who got more than 1 dose of Sertraline.

DR. CASEY: Is the Committee ready to address by vote question number one on efficacy? Or would they like to have more discussion? Dr. Hezel?

DR. HEZEL: I have sort of a practical question to those of you who would prescribe this. Will there be any problem, given that most of the symptoms are GI, that you need to prescribe it with food, which will have relevance, of

course, to efficacy? In other words, if the patients are feeling bad GI-wise, will they be able to take it with food?

DR. LIEBERMAN: I do not know if this is the case but my impression was that this was not an acute effect of administration adverse reaction but was more of a chronic effect that occurred as a result of the sustained pharmacologic actions of the drug. So it is not like lithium where you take it and you may feel a little nauseous.

DR. TAMMINGA: You certainly could never answer the question unless you knew the time course and nature of the nausea.

DR. LIEBERMAN: That is right.

DR. DAVIS: I would like to comment because we had these sheets passed to us but the audience may not have seen them.

DR. CASEY: Please do.

DR. DAVIS: We have been discussing the question of symptomatic volunteers and I gather that a substantial part of these patients were symptomatic volunteers. Of course, a symptomatic volunteers can have severe endogenous depression and they may not be able to afford private care and they may seek care from people offering free care. It raises problems of methodology but it is a complicated question. It is a very difficult question because the field of psychiatry does not know how to characterize who really has genuine endogenous

depression and who does not. Since psychiatry cannot answer that question, it is just an unknown.

But in terms of relative efficacy in the more severely depressed patients in the outpatient trials versus the moderately severely depressed -- maybe it would be better to say in the relatively severe versus the moderate because none of these are severely depressed patients, it looks like the drug-placebo difference is better in the sicker patients among the symptomatic volunteers, which would be reassuring from that point of view. Indeed, in one of these studies the evidence for efficacy is almost nil in the less sick group.

If you look at the final Hamilton's, in the severe group for Sertraline it is 13.4, amitriptyline, 12.3, placebo, 18.6.

DR. CASEY: John, are you reading from table XI or BBI?

DR. DAVIS: I am reading XI.

DR. LEBER: Could we find out what table XI is because it has a peculiar label?

DR. RYDER: I apologize. As I said, you are reading direct SAS output so it is not especially user friendly. Table XI is what was displayed on the screen before. It is data from protocol 104. It is the change from baseline to last visit, LOCF, for the HAM-D total score in the Sertraline, amitriptyline and placebo groups, shown on

the left, and they are categorized according to their baseline HAM-D total severity, the three categories listed, mild, moderate and severe.

DR. LEBER: This is the intent-to-treat group?

DR. RYDER: Yes, it is.

DR. DAVIS: And the other one?

DR. RYDER: That is table XI. Table BBI is the same information from protocol 315. I do not think data were shown. This was the study that did not show efficacy versus placebo. It was the other outpatient study comparing Sertraline, amitriptyline and placebo.

I would be very cautious in reviewing the mild data because the N is extremely small. There were no mild's in protocol 315. That is why the column is absent. In protocol 104 there were 3, 1 and 2.

DR. DAVIS: I would like to go ahead and read the other one too.

DR. CASEY: That is 103?

DR. DAVIS: No. If I understand it right, it is 315 and that is the largely non-United States study in which Sertraline was relatively unimpressive, only slightly better than placebo and substantially less than amitriptyline.

If you look at the more severe group, final Hamilton's for Sertraline are 14.9, for amitriptyline, 16.1 and placebo, 20.8. Changes for Sertraline were minus 13.1,

minus 12.3 and minus 7.4. That looks relatively favorable for Sertraline, being very similar to amitriptyline in the sicker group. The less favorable results were in the milder group.

So there is somewhat of a paradox here where Sertraline was unimpressive in inpatient studies, although I have a question about the inpatient studies as to what concomitant medications were used because there were a lot of concomitant medications.

DR. ESCOBAR: I would agree with John in the case of 104 only. In the case of 315, if you look at the total number of people, those numbers are relatively small. So I guess there is more ammunition to convince me that 104 may be a reasonable trial showing that Sertraline is an effective antidepressant. I am not that convinced about 103.

DR. LEBER: Well, this table for 315 deals with 19 patients across all 3 patients --

DR. ESCOBAR: Yes.

DR. LEBER: It is a subset of a non-randomized comparison. It does not carry much weight one way or the other. We try to look at all the evidence. We try to look at evidence which has some innate persuasiveness to it, whereas, if it looks like a failed trial we treat it like a failed trial. We do not want a lot of failed trials but that happens.

DR. LINNOILA: The reason I was somewhat agitated is that there is clearly tremendous pressure to get new drugs to the market because psychiatric patients are not well served. There are loud complaints from the pharmaceutical companies about how expensive this process is. Then one sees a premiere American pharmaceutical company bringing forward data from clinical studies where experienced academic clinical psychopharmacologists can pretty much a priori say that they are going to be difficulties. You do not start patients at high doses of drugs without titration. It is pretty likely that you are going to get dropouts. On the other hand, you do not put up a multicenter study with 11 different centers and study only 130 patients unless you have months-worth of training, cross-training between the centers so that the ratings are going to be the same.

What agitates me somewhat is that this sloppiness is in the end being charged to the patient in the price of the drugs. That is why I am somewhat agitated.

DR. CASEY: Your opinion was well made. Part of what we are doing, I believe, is mixing up the information we have before us with what we would like to have before us. I agree with many of the comments and there are some parts of the data base that we wish were different but we do not have them. Perhaps we should use some time in giving some guidance to the Division of what we might want to see for the future.

Dr. Leber asked for that, or at least he said he was open to hearing about it.

To be fair to the sponsor, I do not think that we want to change the rules in the middle of the game or near the end of the game. I do not hear anybody saying that but we want to be sure that we are explicit about that.

I still sense that we are approximating coming to an answer on question number one regarding efficacy. I am not sure additional discussion will get us closer. It may get us further away.

We have one issue to consider, and that is the timetable. It is 12:40. Some people might be getting hungry and wish for lunch. The other strategy is to just go forward and see if we can come to some answers. What would the Committee like to do?

DR. LIEBERMAN: I move we go forward.

DR. ESCOBAR: Second.

DR. CASEY: With that proviso, I will give permission to anybody who feels they need to run to the bathroom or to the lunch room, and feel free to do it at any time.

Are we ready for question number one on efficacy?

DR. ESCOBAR: Is the answer only yes or no?

DR. CASEY: No, sir. We could have more discussion. The point is do we need more discussion to address question one on efficacy?

DR. HAMMER: Could the answer be yes or no and also some instructions or requests about labeling?

DR. CASEY: I would like to separate those, that we have an answer about the issue of efficacy and then if we want to provide guidance or an opinion about labeling, I think that would be fine.

DR. HAMMER: Could the answer be yes or no in outpatients only?

DR. CASEY: I think the labeling is something that we should probably leave to the FDA. I would be comfortable with the way it is worded currently with Prozac in terms of its efficacy and then making a specific statement about giving the caveat about inpatient major depressives not being adequately studied. But, again, I think we can address those after the issue of efficacy.

DR. ESCOBAR: The choice for those of us who are not sure is to abstain. Is that right?

DR. CASEY: Abstention is always an option. Yes, you bet. But I do not want to prematurely close the issue. If people want to discuss this or want to have a particular process, let's go ahead and decide what our process should be.

DR. DAVIS: Dr. Leber gave a very good paragraph of what the law was. There were three or four points. The first was what is required for efficacy; what is required for

safety and what is required for labeling. Please repeat what you said before.

DR. LEBER: Basically, the law -- this is not a regulation but the law that carries it -- says that when an application is submitted the FDA shall review it and, within that statutory time limit that is never met, reach a conclusion and approve the drug unless it finds from that review that the drug is unsafe for use under the conditions enumerated in the labeling or that there had been a lack of test to show that the drug is safe under those conditions. That is, it shall approve it unless it finds there is a lack of substantial evidence.

Then it defines what substantial evidence is, and I will sort of paraphrase it, that is, evidence derived from adequate and well-controlled investigations, including clinical investigations, that would allow experts qualified by training and experience to reach the conclusion or reach the judgment that the drug will have the effect claimed for it in the labeling. It is not, as they say, a preponderance of evidence standard but a substantial evidence standard, and you can all get your attorneys to help you will that one.

Finally, we are to approve a drug unless we find that its labeling is false and misleading in some particular. For example, you might be able to impose a restriction on the indication because you would say that to say it is an

sgg

antidepressant without mentioning that it fails to have an effect demonstrated in inpatients would be misleading. So there are ways you can get around it.

Basically that is what we are asking you. We are not asking you about the preclinical tox. We are not asking you about chemistry and so on.

DR. DAVIS: There is a problem there because we have been asked to comment on labeling without knowing the labeling.

DR. LEBER: Well, that has been a tradition, having read your sense of what you are worried about, to negotiate the specifics with the firm. The reason for that is that we are trying to maintain a certain degree of equity in dealing with all of the regulated industry. I will tell you candidly that on an antidepressant, for example, we try to state that the drug is approved for the treatment of depression. Then we put the qualifiers after that. That stems from our interest in avoiding what I call "pseudo-specific claims" -- depression in bored housewives, depression in the elderly, depression in those with heart disease, which would elaborate from that.

So, basically, the question we want to know from you is, given the evidence that you have seen on efficacy that has been submitted, do you think it is enough to conclude that there is a lack of substantial evidence of

efficacy?

DR. DAVIS: I think I would infer from that that we should discuss labeling to have some guidelines recorded for the minutes because it may be an issue.

DR. HAMMER: And we have discussed labeling in the past and made recommendations in the past. I just do not remember whether we did that before or after approval for efficacy.

DR. CASEY: I believe it was after.

DR. LEBER: There is no point having labeling if you do not have an approved drug. Efficacy is the sine qua non for approval.

DR. DAVIS: In terms of general discussion before we get down to brass tacks, one thing that was suggested before, which I would like further clarification on, is this matter of when you have multiple studies -- one could do a number of studies and then pick the study which was most efficacious and go with that, which would introduce a statistical bias. What is the law in terms --

DR. LEBER: The law, as far as I know, never discussed the issue of multiplicity. However, I think the interpretation locally would be that if you did a hundred studies and found five that were effective, something would be wrong with the error rate and we would not approve it.

-- has not been unusual in the past to have many

failed studies and yet have the drug approved. I have talked about this and so have other members of the staff, we do not have a systematic program for a meta-analysis within an NDA. Maybe there ought to be. Jerry Levine, for example, is very interested in types of review of the data elements that one would look at in reaching a conclusion. If this were a series of trials with a common endpoint, maybe we would use a Yussef-like (phonetic) method. But we do not. That probably has to do with the fact that the law has said "substantial" rather than "preponderance" and you have to believe that there is enough evidence here to allow experts to conclude that the drug will have the effects claimed for it. The drug claim will basically be that it is an effective antidepressant with this proviso that we do not know if it works in in-patients, or something to that effect. Tom, why don't you have a whack at expanding on this?

DR. LAUGHREN: I agree. I would prefer to have an initial opinion about efficacy prior to going on to talking about advice for labeling. I think that is where you have to start.

DR. LIEBERMAN: I think we have gone as far as we can in terms of examining the evidence available to us. I move that we take a poll on efficacy and then move to what our attention seems to be focusing on, which is how to qualify the recommendation for efficacy.

DR. HEZEL: I have a point for clarification. Several Committee members have questioned protocol 103. So 104 is 489 patients and we have some consensus on our Committee for efficacy?

DR. CASEY: Are you asking to clarify the number of patients in study 104?

DR. HEZEL: That is right.

DR. CASEY: Does anybody know immediately?

(Dr. Choudhury shows Dr. Hezel the documentation)

DR. CASEY: Dr. Hezel, have you had that question answered satisfactorily? It is more than 400, less than 500, depending on how you work with the numbers.

DR. LEBER: We can get you an exact number but it may change.

DR. CASEY: I sense that we are moving closer to having a vote on question one. For those who are wishing that somebody would give an opinion, I will state that I think that there is sufficient evidence for efficacy, though I hope that the lesson from today is that additional work will be well rewarded when it is put into the front end of studies and that, as the years go by, we will see more and more advancement in the quality of psychopharmacological research brought to the Committee.

Shall we take a vote on question number one? The issue before us is has the sponsor provided evidence from

more than one adequate and well-controlled clinical investigation that supports the conclusion that Sertraline is effective for the treatment of depression? Those who believe the sponsor has provided such information, please vote now.

(Show of hands)

Six in favor, including myself. Those who believe the sponsor has not provided adequate evidence, please raise your hands.

One. Those wishing to abstain, please do so.

(Show of hands)

Two. Six yes, two abstaining and one no.

We will move on to the second issue? Has the sponsor provided evidence that Sertraline is safe when used in the treatment of depression? Those in favor, raise their hands.

(Show of hands)

Nine. That is unanimous.

Now to the issue of labeling. Comments?

DR. LEBER: It would be useful for you to say affirmatively now what you would like to see. The exact wording has to be negotiated but it would be nice to know what your concerns are in an organized way so that when we come to write labeling we can carry forward your wishes.

DR. LINNOILA: My concern is that the sponsor has not provided us with any data which would suggest efficacy in

the treatment of depressed inpatients.

DR. HAMMER: I agree. I think the sponsor has failed to show efficacy in inpatients. I would like the labeling to say something that communicates that.

DR. PRIEN: Have we separated the issue of severity from inpatient/outpatient status?

DR. TAMMINGA: We have not done that fully. In the outpatient studies the Company has separated the people into mild -- although there are not enough of them -- moderate and severe. If one were looking at inpatients, would the severity criteria be different? I am not sure how the inpatient/outpatient question is different.

DR. DAVIS: I think one has to think of this in the context of symptomatic volunteers. We have data in the symptomatic volunteers classified as mild, moderate and severe but those are still severe symptomatic volunteers. So we do not know about real depressed outpatients, although I am sure many of these are. There are somewhat inconclusive inpatient trials, which makes it a complicated and difficult matter to write labeling for.

DR. CASEY: The labeling may state just that, that issues of severity for inpatient and outpatients are not yet clarified.

DR. HAMMER: But I think it would be pretty plain [illegible] demonstrate efficacy in

inpatients. That is four or five words. You cannot get much plainer than that.

DR. LEBER: Except that it can have less meaning than that. See, that is my concern. Again, I want to emphasize to all of you that I do not think, Dr. Linnoila's words not withstanding, that anyone really knows what it means to say someone is an inpatient. In fact, what we need to do in some way, and I think this is a point John Davis was getting at, we do not really have independent, objective measures of severity of illness in depression. These are all rating scales made by people.

So it is conceivable, to show you a dishonest way out of this, that a sponsor could arrange to have hospitalized the kind of patient that, on the basis of the data they now know, is likely to respond. Then they would have met your criteria of showing an effect in inpatients and have totally violated the spirit of what you want to check, whether the severe, endogenomorphically depressed, perhaps highly suicidal, dangerous, difficult to manage patients will respond to this treatment in an experiment in which they also respond to something like amitriptyline or even ECT. Don't you?

DR. DAVIS: I would say that two inpatient studies were done, with relatively small sample sizes, with some concomitant medications, which did not show efficacy, and let

the readers of the PDR draw their own conclusions.

DR. LAUGHREN: At some point it becomes an issue of how much you can put in labeling. I gave you the example from Prozac where we were in a very similar situation and did not have a positive study in depressed inpatients. To go to the length of describing exactly what was done would add to the length of labeling. I am not sure that it would add enough additional valuable information to justify it. I think that is my concern.

Obviously the issue here is that there is a disagreement about what inpatient status conveys even on this Committee.

DR. DAVIS: I am not sure there is a disagreement. I agree entirely that it is somewhat indeterminate. I agree with what Dr. Leber said, there are problems with inpatient studies as well as outpatient studies and it is hard to have the exact balance.

One thing I do like about FDA labeling is your tables where you will often have a table of side effects comparing a new drug with placebo so the reader can read the numbers. I think you could do something similar here with a carefully chosen sentence.

I would like to give you some discussion, at least from me on this Committee, to allow you support for doing [illegible] if you would want to do it. I also agree with what you

said, that in labeling you cannot have every possible qualification and you have to try and put it in well chosen sentences to get some sort of essence.

DR. LIEBERMAN: I have one informational question and then a comment as to labeling. It is sort of late to be raising this but in terms of the severity issue, do we know what proportion of the major depressive disorder patients in the sample may have been of the melancholic subtype? Because of severity and in terms of how it relates to inpatient versus outpatient patients and symptomatic volunteers that might be useful to know.

The comment as to labeling has to do with the dosage. The dosage that is being recommended for utilization is 50-200 mg. So if we are going to comment as to the lack of information determining efficacy in more severe forms of depression, we might qualify that to say in the dose range that is currently being prescribed.

DR. CASEY: The sponsor says that they do have the information available, if you feel that it is contributory at this point, regarding subtype diagnoses.

DR. LIEBERMAN: I would be interested in hearing it, yes.

DR. RYDER: As I think Dr. Lee mentioned, patients were categorized according to DSM-III and 62 percent were major recurrent; 34 percent single episode and for melancholy

and non-melancholy subtypes, if I remember correctly, the distribution was about half and half.

DR. LIEBERMAN: I, personally, find it somewhat reassuring to know that a significant proportion of the sample was comprised of melancholic subtype patients.

DR. DAVIS: It is quite possible that further information from the sponsor could clarify this matter because one could get further information about drug/placebo efficacy in their data base classified by criteria like endogenomorphic versus non-endogenomorphic drug/placebo effects. If the drug/placebo effects are larger in the more carefully diagnosed depressed patients, it would allow stronger labeling in terms of specificity for the depressed disease.

DR. CASEY: It may be just as problematic as those words also mean something special to each one of us but we may not yet have cross-validated definitions of what those are.

DR. LEBER: I am going to give you a challenge. I mean we start looking at subcategories of this great unwashed mass called depressed patients and the implication is that knowing the subclassification in some way predicts treatment response, course, outcome or the like. I would like to see the evidence before we jump to those conclusions that tells [illegible] in any way tell you what the

treatment response is going to be. If you do not have that, it is like saying that 60 percent of the sample is old ladies from Queens, and that is not going to tell you anything.

What we need to find out about are the markers for treatment response. I think they are going to be biological eventually. Certainly Dr. Lieberman would agree with that concept. When we have those, then we can do it but right now most of the clinical descriptors do not tell us anything, including inpatient status I think.

DR. LIEBERMAN: You are right in the sense that there are imperfect and imperfectly validated sub-syndromes. But, on the other hand, there is significant evidence in terms of biologic measures, as well as placebo response rates, which are associated with melancholia as a subtype as opposed to the larger category of MDD.

DR. LEBER: I would like to see the meta-analysis on that.

DR. CASEY: I would like to bring up an additional point that we have not directly touched on. That is, whether drugs in this class lead to patients attempting or successfully committing suicide, as has been suggested about another compound in this class. I read the data in the presentation today as there not being evidence of that for Sertraline. Are there additional comments? Did other people read this similarly or the same? Everyone is nodding their heads that

yes, they read it the same as I do. There is not evidence that Sertraline produces increased risk of attempted or successful suicide in the patients reported.

DR. ESCOBAR: On a short-term basis.

DR. CASEY: Within the limits of the data that we have had presented today.

DR. TAMMINGA: That is for 44 weeks.

DR. HAMMER: And no more so than successful treatment of extremely depressed patients can increase the risk of suicide by enabling them to do it.

DR. LEBER: While we are at it, it shows you how times change. When we first looked at fluoxetine, the great sword of Damocles that hung over us was the issue of zimelidine with ascending paralysis and its flu-like syndrome. Now with several years under our belt, I guess we have accepted one or two cases and we do not think it is higher than the underlying risk, and that has gone to sleep. Now with fluoxetine being pegged on the issue that was just brought up for possibly inducing the strange subset of ideational behavior about suicide, that is why this question comes up.

This seems to be remarkably clean as antidepressants or, in fact, all drugs go for its safety data base. How many seizures did we have? I do not think it was even mentioned.

DR. LAUGHREN: I believe there was one seizure but

stopped taking his anticonvulsant.

DR. RYDER: That is correct. If I remember correctly, that patient was taking 100 mg of Sertraline. The drug was never stopped. The patient had sub-therapeutic levels of his maintenance anticonvulsant. The anticonvulsant was reestablished and Sertraline was continued for an additional -- if my memory is right -- six weeks, until the end of the study, without any further episodes.

DR. LEBER: The point that I was [illegible] to get some concurrence with you on is that complaints were made for many years that the second generation of antidepressants did not work and did not do anything. But we are now beginning to get a class of second and third generation antidepressants which work. You have your concerns about how much but they seem to be presenting with a very different panoply of side effects. There are no cardiovascular effects to speak of. There are no anticholinergic effects to speak of -- relatively free of some of the things that have troubled us with some of the tricyclics in the males. Maybe the last three or four have been like that. So there are dimensional tradeoffs.

Anybody disagree with that assessment? I am not trying to make a commercial but for those who are neither in the drug industry nor in the academic community who may read our transcripts.

DR. HAMMER: There is no such thing as a side

effect free drug. Even placebo.

DR. LEBER: There is no free lunch, right.

DR. DAVIS: I would agree with you. I think some of these newer drugs have significantly less side effects.

DR. CASEY: Are there other issues that the Committee members would like to put on the table for discussion?

DR. DAVIS: I would like to comment the dose-response study. Had there been a nice linear dose-response curve, everybody would have been enthusiastic. There may be a hint of a flattening out but that is important information. It is important to see things in perspective. I would see some of the design things as a step forward.

DR. CASEY: Agreed.

DR. DAVIS: I also would like, from a general philosophic point of view, to follow up on Dr. Laughren's comments. If one is too conservative, one does not release drugs soon enough to the market and one cannot have answers on every single question. But I thought Dr. Laughren's list of questions for where there should be some data was a good list. I think that should be taken seriously. But you cannot answer every single possible thing with a very large sample size study. But it would be good to have some reasonable information, hopefully, some quantitative and some

DR. CASEY: I agree. They are very reasonable. They are the kinds of things you would like to have for practical use of how you use a drug with a patient or group of patients.

DR. DAVIS: It was too good a list of information not to speak in support of.

DR. LAUGHREN: Part of my reason for presenting it was in hope of getting companies to think about these things early on in development. I think many of those questions could have been answered without making much in the way of a change in the trials that were already being done, and may have been answered quite easily if they had been thought about early enough.

DR. CASEY: We are coming to a close but, first, Dr. Davis has one more comment.

DR. DAVIS: Speaking in terms of perspective, I think it is very important to recognize that the study was very brief in duration and I agree with Dr. Lee's comments, but that, at least to me, was important evidence for efficacy within the limitations of the study. That is another new positive development.

DR. CASEY: Yes, though there were many limitations.

DR. PRIEN: One other point that was brought up, I think we should continue to encourage companies, the FDA and academia to pay more attention to continuation and longer-

term maintenance studies, not only for antidepressants but other psychotropic drugs as well, and perhaps work out a design paradigm that is acceptable to everyone.

DR. CASEY: Yes. One way to put the message is not to just do what somebody else did before, what they have always historically done, but put some thought into this and expect that you will be asked to advance the field step by step. Dr. Leber?

DR. LEBER: I have one other thought about long-term maintenance studies. You all may be aware that Stewart Montgomery, of the CSM, has gone around saying that the requirement in the world's new economic order in Europe, I guess, will be that there be a maintenance study of patients who have recovered. He claims to have done it with fluoxetine against placebo. I think he took patients who were six months into their recovery and re-randomized them to placebo and fluoxetine. I do not remember the full design but he then carried them forward. That was to show maintenance against relapse.

It is the kind of thing that I think we would like to do if we can get people to do it. The question again is regulatory authority and is that within the set of minimum demands you can make?

DR. LINNOILA: I think the other message would be [illegible] should design their multicenter studies

sensibly. If one takes a very large number of centers with very few patients in each center, it becomes an impossible task to have uniform diagnostic criteria and uniform ratings. That can lead to the kind of data that we have looked at here. From the marketing point of view, it is perhaps great to have many people have some experience with the drug. It is a quick launch. But it is not good for proving efficacy.

DR. CASEY: Agreed. Dr. Davis?

DR. DAVIS: One of the most important points in treating a recurrently depressed patient is use of long-term drugs to prevent relapse. Within the limitations of this particular study, there was a fair drug/placebo difference in the early periods. It may be something that would be important to handle regulatorily in thinking about drugs in the future as to what is said in the package insert about whether a drug company did or did not prove efficacy for maintenance. In other words, if the law is not changed or if the law does not address acute versus long-term effects specifically, one would have to handle it in what the package insert says about long-term maintenance treatment as opposed to acute treatment. I am speaking to support Dr. Prien's remarks that it is an important side of the whole picture.

DR. CASEY: Certainly the theme that has been evolving through our Committee over the past several sessions is the importance of seeing how these drugs are studied in

ways in which they are actually used.

DR. HAMMER: I would like to make one more comment on the labeling. I think that if the labeling for this drug is identical to Prozac, and I am not going to comment on whether I think that is appropriate for Prozac, but if it is identical, that would allow some readers to draw the inference that had the studies been done properly the drug would have been shown effective. Whereas, I think that if you just make a plain statement that they failed to show efficacy in inpatients, that will not encourage the reader to make the inference that had the studies been done correctly the drug would have been shown to be efficacious.

DR. CASEY: I think your point has been taken by the Agency.

Once earlier I said we are getting to a close. I think we are getting to closure. I want to thank the Agency for their very clear, concise and well-thought out presentations. I want to thank the sponsor for their clear and well-thought out presentation and their availability to provide information when we asked for it. I want to thank Mr. Bernstein for his work in making this meeting run well and, finally, to thank the Committee for all their efforts. The meeting is adjourned.

(Whereupon, at 1:20 p.m., the Committee adjourned)

C-E-R-T-I-F-I-C-A-T-E

I, D. Gavrisheff , the Official Court Reporter for Miller Reporting Company, Inc., hereby certify that I recorded the foregoing proceedings; that the proceedings have been reduced to typewriting by me, or under my direction and that the foregoing transcript is a correct and accurate record of the proceedings to the best of my knowledge, ability and belief.

/s/ D. Gavrisheff