HALT-C Statistical Analysis Plan

A. INTRODUCTION

The HALT-C trial is a randomized controlled clinical trial in which 810 patients will be randomized to treatment or control and followed for 3.5 years at 9 clinical centers. Details of the study design are included in the protocol and are briefly summarized below. The purpose of this document is to provide a detailed analysis plan for the trial. This draft of the document includes the preliminary analysis plan. The analysis plan will be finalized during the lead-in period, prior to the randomization of the first patient.

B. PRIMARY AIM OF THE HALT-C TRIAL

The primary aim of the HALT-C trial is to determine if, in patients with chronic hepatitis C who failed previous interferon therapy, 3.5 years of interferon therapy will

1. In patients with fibrosis only, prevent progression of advanced fibrosis to cirrhosis and
2. Reduce the risk of hepatic decompensation, the need for hepatic transplantation, the risk of developing hepatocellular carcinoma (HCC).

C. SUMMARY OF TRIAL DESIGN

Approximately 1,200 patients 18 years of age or older who meet the inclusion and exclusion criteria as defined in the protocol will be entered into a lead-in phase. They will be treated with a combination of Peginterferon alfa-2a and ribavirin for a period of 24 weeks. Patients who have no detectable HCV RNA at week 20 will continue on combination therapy until week 48.

Patients who do not clear virus will be randomized 50:50 at week 24 to receive either Peginterferon alfa-2a alone or no further therapy for three and a half years. Randomization will be stratified by clinical center and by the presence of fibrosis or cirrhosis (baseline histology Ishak fibrosis score of 3-4 versus 5-6). Both randomized groups will be monitored quarterly during these 42 months. Repeat biopsies will be obtained at 24 and 48 months after the start of the lead-in phase.

The primary outcome variable of the trial is a composite variable consisting of any of the following:

1. An increase in the Ishak fibrosis score by 2 points or more at the year 2 or year 4 biopsies;
2. Death from any cause;
3. Development of HCC;
4. A Child-Turcott-Pugh (CTP) score of 7 or higher at two consecutive visits;
5. Variceal hemorrhage;
6. Ascites;
7. Spontaneous bacterial peritonitis; or
8. Hepatic encphalopathy.

Secondary outcome variables include the following:

1. Quality of life;
2. Serious adverse events;

3. Events requiring dose reductions (in both groups); and
4. Changes in Ishak fibrosis scores from baseline at the year 2 or year 4 biopsies.

The sample size of 810 patients (90 at each clinical center) was based on the following assumptions:

1. A two-sided alpha (type I error) of 5% and a power (1 - type II error) of 90%;
2. Equal allocation of patients to treatment and control;
3. The statistical test will be a chi-square test of a dichotomous outcome (event/no event) (Friedman, 1996);
4. An annual event rate in the control arm of 6% per year;
5. An annual event rate in the treatment arm of 3% per year (50% reduction); and
6. Noncompliance rates of 5% per year in the control arm and 3% per year in the treatment arm in years 2 to 4, resulting in 3.5 year rates of 18.7% in the control arm and 10.6% in the treatment arm.

D. ANALYSIS OF BASELINE DATA

For this trial, baseline data are collected prior to the lead-in phase. Some additional data are collected at week 20 prior to randomization. Two sets of baseline analyses will be performed:

1. Lead-in patients and
2. Randomized patients.

The analyses of baseline data will include clinical center comparisons, comparisons of patients in the two strata (fibrosis or cirrhosis), and comparison of the two randomized treatment groups.

Baseline characteristics that form continuous variables will be compared using analysis of variance. Data will be transformed (e.g. logarithms) prior to analysis if transformation is needed to meet the assumptions of normality and homogeneity of variances that underlie these methods. Analyses will be performed using SAS Proc GLM (SAS Institute Inc., 1989, 1997) with stratum and clinical center as class variables. A weighted analysis will be performed using Type II sums of squares (Lin 1997; Senn 1998).

Non-parametric methods, such as the Wilcoxon-Mann-Whitney test, will be used if data fail to meet these assumptions even after transformation (Lehman, 1975). These methods will also be used for the analysis of ordinal categorical variables. The analyses will be stratified by fibrosis/cirrhosis and by clinical center.

Binary variables will be analyzed using the chi-square test. The Cochran-Mantel-Haenszel statistic will be used to adjust for strata and clinical centers (Agresti, 1990; Fleiss, 1981).

The comparison of randomized groups at baseline has been criticized (Senn, 1992).
It is expected, due to randomization, that the patients in the treatment arm will, on average, be similar with respect to variables that might influence outcome. However, for imbalances that persist despite randomization, co-variates or strata will be included in the analysis of the primary and secondary endpoints.

Baseline analyses will also examine the extent and distribution of missing data.

E. ANALYSES OF LEAD-IN PHASE

The primary outcome for the lead-in phase is the disappearance of HCV RNA. Logistic regression will be used to determine the characteristics that are predictive of a response to Peginterferon alfa-2a and ribavirin. Covariates to be included in these analyses include the following;

- Age
- Gender
- Race
- Baseline fibrosis or cirrhosis
- History of alcohol abuse
- HCV genotype
- Years since exposure (if known)

F. ANALYSES OF THE PRIMARY OUTCOME

The primary analysis will be an intent-to-treat analysis. All randomized patients will be included. This includes patients who are later found to be ineligible and patients who do not receive the assigned treatment or become noncompliant during the course of the trial. A two-sided significance level of 0.05 will be used for all analyses.

The analysis of the primary outcome is complicated because this outcome is a composite outcome that consists of both the results of biopsies obtained at two-year intervals and events that may occur at any time during the follow-up period. In addition, the biopsy outcome, the development of cirrhosis, can occur only in patients in stratum 1 (fibrosis) and clinical events will occur primarily in patients in stratum 2 (cirrhosis).

We propose the use of a nonparametric test based on generalizations of the Wilcoxon rank sum test (Finkelstein and Schoenfeld, 1999; Gehan, 1965; Peto and Peto, 1972; Tarone and Ware, 1977). For stratum 2 (cirrhosis) the test will be based on the time-to-event data. For stratum 1 (fibrosis) the test will be based on the biopsy results at 24 and 48 months. Analyses will be stratified by clinical center and a combined test will be conducted as described by Finkelstein and Schoenfeld, 1999.

For stratum 1, this outcome variable depends on the availability of follow-up biopsies in most patients. If biopsies are not obtained because the patient has developed symptoms of liver disease, then it will be assumed that the biopsy would have shown cirrhosis. If the biopsy was not obtained and the patient did not experience any clinical events, then the result will be missing. The effect of these missing data will be assessed by performing additional analyses in which the results are assumed to be either cirrhosis absent or cirrhosis present. Patients with missing biopsy results will be compared to other patients in an attempt to identify baseline characteristics that would predict missing results.

Alternative methods of analysis will be investigated prior to the start of patient randomization. These methods include interval-censored survival data (Finkelstein, 1986; Lindsay and Ryan, 1998; Sun, 1997) and methods for the analysis of recurrent events (Ghosh D, 2000; Kelly and Lim, 2000; Mahe and Chevret, 1999; Zeger and Liang, 1986).

G. INTERIM ANALYSES

An interim analysis will be performed once 50% of the Year 2 biopsies have been evaluated by the Central Pathology Committee. We recommend implementation of an early stopping rule based on an O'Brien-Fleming group sequential plan (O'Brien, 1979). This will be implemented using the alpha spending approach of Lan and DeMets (DeMets and Lan, 1994; Lan and DeMets, 1983).

Once the early stopping rule is established, the DCC will produce interim analyses for the DSMB that will include:

1. Standard recruitment graphs comparing actual to goal over time and projected completion point.
2. Detailed tables and figures of recruitment experience at each CC (Number, by month or quarter by CC, as well as gender, age and minority cross-tabulations by CC).
3. Tables comparing baseline characteristics, cumulatively, at each interim analysis by (blindly labeled) treatment groups. These tables will be presented in aggregate as well as by CC.
4. Protocol and eligibility violations in trial patients will be reported since the last interim analysis and cumulatively, by CC with full documentation.

It is anticipated that the proposed stopping rule will have asymmetric boundaries to reflect the overriding concern with minimizing the possible deleterious effect of peginterferon alfa-2a (if observed). We also strongly advocate maintaining, as long as feasible, blindness of the DSMB to actual treatment group, by arbitrary labels such as A and B. This provides additional security if copies of reports are seen by others and may assist the DSMB members in maintaining objectivity.

H. ANALYSES OF SECONDARY ENDPOINTS

Comparisons of the two arms at four years will be conducted as described for the baseline data. Longitudinal methods will be used to analyze changes in quality of life and other measured variables (Diggle et al., 1994). These will include the calculation of slopes where the change is expected to be linear or calculation of area under the curve for nonlinear changes.

REFERENCES

Agresti A. Analysis of Categorical Data. New York: John Wiley 1990.

Cox DR. Regression models and lifetables. Journal of the Royal Statistical Society 1972; 34:187-220.

DeMets L and Lan KKG. Interim analysis: the alpha spending function approach. Statistics in Medicine 1994; 13:1341-1352.

Diggle PJ, Kiang KY, and Zeger SL. Analysis of Longitudinal Data. Oxford University Press, New York, NY, 1994.

Finkelsein DM. A proportional hazards model for interval-censored failure time data. Biometrics 1986; 42:845-854.

Finkelstein DM and Schoenfeld D. Combining mortality and longitudinal measures in clinical trials. Statistics in Medicine 1999; 18:1341-1354.

Fleiss JL. Statistical Methods for Rates and Proportions. New York: John Wiley 1981

Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials, 3rd ed., Mosby-Year Book , Inc. St. Louis, MO, 1996.

Gehan E. A generalized Wilcoxon test for comparing arbitrarily single censored samples. Biometrika 1965; 52:203-223.

Gosh D. Methods for analysis of multiple events in the presence of death. Controlled Clinical Trials 2000; 21:115-126.

Kelly PJ and Lim LL-Y. Survival analysis for recurrent event data: an application to childhood infectious disease. Statistics in Medicine 2000 19:13-33.

Lan KKG and DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983; 70:659-663.

Lehmann EL. Nonparametrics: Statistical Methods Based on Ranks. New York, NY: McGraw-Hill 1975.

Lin Z. An issue of statistical analysis in controlled multi-centre studies: How shall we weight the centers? Statistics in Medicine 1999; 18:365-373.

Lindsay JC and Ryan LM. Tutorial in biostatistics: Methods for interval-censored data. Statistics in Medicine 1998; 17:219-238.

Mahe C and Chevret S. Estimation of the treatment effect in a clinical trial when recurrent events define the endpoint. Statistics in Medicine 1999; 18:1821-1829.

O'Brien PC and Fleming TR. A multiple testing procedure for clinical trials. Biometrics, 1979; 35:5490566.

Peto R and Peto J. Asymptotically efficient rank invariant test procedures. Journal of the Royal Statistical Society Series A 1972; 135:185-207.

SAS Institute Inc., SAS/STAT® User's Guide, Version 6, Fourth Edition, Volumes 1 and 2, Cary, NC: SAS Institute, Inc., 1989.

SAS Institute Inc., SAS/STAT® Software: Changes and Enhancements through release 6.12, Cary, NC: SAS Institute, Inc., 1997.

Senn S. Testing for baseline balance in clinical trials. Statistics in Medicine 1994; 13:1715-1726.

Senn S. Some controversies in planning and analysing multi-centre trials. Statistics in Medicine 1998; 17:1753-1765, 1998.

Sun J. Regression analysis of interval-censored failure time data. Statistics in Medicine 1997; 16:497-504.

Tarone R and Ware J. On distribution-free tests for the equality of survival distributions. Biometrika 1977; 64:156-160.

Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. Journal of the Royal Statistical Society, Series B 1976; 38:290-295.

Zeger SL and Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986; 42:121-130.