S. Tan
January 8, 2008

Dataset Integrity Check for the Study of Viral Resistance to Antiviral Therapy for Chronic Hepatitis C (Virahep-C) Baseline Files

Note. This current version, dated 01/08/08, replaces the previously released DSIC dated 12/17/07.

The Virahep-C multicenter treatment trial was designed to assess the rates of response to peginterferon combination therapy for hepatitis C among African-American (AA) and Caucasian-American (CA) treatment-naïve patients. The goal of the study was to elucidate the mechanisms of antiviral resistance among patients who fail to respond to peginterferon combination therapy. As a partial check of the integrity of the Virahep-C main study datasets archived in the NIDDK data repository, a dataset integrity check (DSIC) was performed to verify that selected published results from the main study can be reproduced using the archived datasets. The DSIC consists of a small number of analyses performed to duplicate published results reported by the Virahep-C Study Group [1] in *Gastroenterology* in August, 2006. Results of the DSIC are described below.

The intent of this DSIC is to provide confidence that the data distributed by the NIDDK repository is a true copy of the study data. Our intent is ***not*** to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected on a first (or second) exercise in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study data coordinating center (DCC), however this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy that is observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, *unless staff of the NIDDK Repository suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff*. We do, however, document in footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

Dataset Integrity Check. The DCC submitted 34 data files representing the raw data collected for the Virahep-C main protocol, together with corresponding data collection forms, brief dataset descriptions, variable labels, and SAS formats (value labels). Contents of the archived raw datasets exactly matched the dataset descriptions provided by the DCC. Missing data values were qualified, using SAS formats, as “Unknown”, “N/A”, “Refused”, or “Missing”. Analysis datasets used for results published by the Virahep-C Study Group were not provided at the time of this DSIC. A representative of the DCC was available for general questions regarding study and data protocols; however, reconstruction of the analysis dataset was left primarily to Repository analysts. For purposes of the DSIC, an analysis dataset was recreated by Repository staff from portions of the 34 raw data files. Results of analyses on the reconstructed analysis dataset were compared to published results. A percentage of published results were replicated, selected to assure the quality of the archived datasets.

Analysis Methods. As indicated in the publication, baseline demographic and clinical characteristics among racial groups (AA and CA) were compared using the $?^2$ test of association with continuity correction or exact tests. The nonparametric Wilcoxon rank-sum test was used to test for racial differences in the distributions of continuous variables. The primary end point, sustained virologic response (SVR), was defined as the absence of detectable HCV RNA in serum 24 weeks after therapy was stopped[1]. Accordingly, an outcome indicator denoting presence or absence of SVR was constructed. The DCC provided the following protocol for the calculation of SVR:

> **Definition of Sustained Viral Response (SVR)[2]**
>
> Qualitative tests were completed in duplicate. SVR was defined by qualitative test results at 24 weeks post treatment with a minimum time of 20 weeks post treatment. If both qualitative tests were positive, then participant was considered to be a non-responder (NR) for SVR.
>
> If qualitative results were not available at 24 weeks post treatment:
> 1.) If no further tests were available, the participant was considered to be a non-responder (NR)
> 2.) If subsequent qualitative tests were available then these results were used to determine SVR status.

Associations between measures of interest and SVR were reported as relative risks with 95% confidence intervals. Estimates of relative risk were assessed using generalized estimating equations[3]. Statistical analyses were performed using *PC-SAS v.9.1 (Cary, NC)*[4].

Analysis Cohort. The study publication reports that of the 401 patients who met the entry criteria and subsequently enrolled, 196 were AA and 205 were CA. Upon examination of the raw datasets, Repository analysts constructed an analysis dataset as follows:

1.) Patients with baseline lab or viral load data were identified (N=403)
2.) Patients who withdrew consent at baseline visit (N=1)
3.) Patients who were found to be ineligible during the course of the study were removed (N=1)[5]
4.) The analysis dataset was limited to the 401 remaining patients, 196 AA and 205 CA.

Baseline Comparisons. The following results, reported in the study publication (text, p.472 and table, p.473), were replicated by analyses of archived data for the 401 patients. AA and CA did not differ significantly in regards to sex and age. Education greater than high school level was less common among AA patients. AA patients were heavier and more likely to have a history of diabetes and hypertension. AA patients were more likely to be currently smoking than CA patients, but the difference is not significant. AA patients had lower levels of serum ALT, albumin, total bilirubin, and hemoglobin, compared to CA patients. AA patients also had borderline lower levels of serum AST and WBC compared to CA. HCV-RNA levels were similar between the two racial groups[6]. Actual medians and proportions for these baseline characteristics differed slightly when comparing archived data results to published results (DSIC Table 1).

---

[1] Per study publication, p. 471.

[2] Per E-Mail communication with the DCC, 8/13/07.

[3] For purposes of the DSIC, generalized estimating equations (GEE) were used to estimate relative risk of univariable associations with SVR. (The study publication does not explicitly state the statistical methods used to assess univariable associations.)

[4] The study publication states that *SAS v.8.02* and the R language and environment were used in analyses.

[5] Patient #**295076** was found to be ineligible, even though the patient collected data for baseline and treatment visits.

[6] Communication with the DCC verified that published baseline viral load levels were in units of **$\log_{10}$**. These units are not stated anywhere in Table 1 of the publication (p.473). By assessing $\log_{10}$ of the archived intent-to-treat viral load variable <*vload_itt*>, we succeeded in replicating published baseline values.

DSIC Table 1. Selected[†] Baseline Characteristics of Virahep-C Participants.

[†] *analyses selected from publication Table 1, p. 473.*

Characteristic	Source* Dataset / variable	African Americans						Caucasian Americans						P-value	
		Archived (n=196)		Published (n=196)		Difference (0)		Archived (n=205)		Published (n=205)		Difference (0)		Arch-ived	Pub-lished
Male, n (%)	SE / sex	127	(64.8)	127	(64.8)	0	(0)	134	(65.4)	134	(65.4)	0	(0)	0.99	0.99
Age, y	SE / (various)	49.2	(45.3, 53.0)	49.0	(45.0, 52.5)	-0.2	(-0.3, -0.5)	48.2	(42.9, 52.4)	48.0	(43.0, 52.0)	-0.2	(-0.1, -0.4)	0.07	0.08
Education	SQ / educ	(n=192)		(n=191)		(-1)		(n=204)		(n=202)		(2)			
>High School		104	(54.2)	104	(54.5)	0	(0.3)	132	(65.0)	132	(65.4)	0	(0.4)	0.04	0.04
≤High School		88	(45.8)	87	(45.6)	-1	(-0.2)	71	(35.0)	70	(34.7)	-1	(-0.3)		
Weight (kg)	SE / wgtlb	(n=193)		(n=194)		(1)								0.001	0.001
		87.3	(78.2, 100.9)	87.5	(78.6, 100.9)	0.2	(0.4, 0)	82.3	(72.7, 95.9)	82.3	(72.7, 95.5)	0.0	(0, -0.4)		
Body mass index, kg/m^2	SE / (various)	(n=191)		(n=193)		(2)		(n=203)		(n=203)		(0)		0.0003	0.0003
		29.5	(26.3, 34.2)	29.3	(26.4, 34.0)	-0.2	(0.1, -0.2)	27.6	(24.4, 31.5)	27.6	(24.4, 31.5)	0.0	(0, 0)		
History of diabetes, n(%)	SE / mxdiab	30	(15.3)	30	(15.3)	0	(0)	9	(4.4)	9	(4.4)	0	(0)	0.0004	0.0004
History of hypertension	SE / mxbp	85	(43.4)	84	(42.9)	-1	(-0.5)	43	(21.0)	43	(21.0)	0	(0)	<0.0001	<0.0001
Currently smoking, n(%)	SQ / cursmk, evsmk	78	(41.1)	78	(41.3)	0	(0.2)	70	(34.5)	70	(34.5)	0	(0)	0.22	0.20
Alanine transaminase level, IU/L	LE / ALT	59.0	(40.0, 90.0)	59.0	(40.0, 88.0)	0	(0, -2)	74.0	(51.0, 138.0)	74.0	(51.0, 138.0)	0.0	(0, 0)	<0.0001	<0.0001
Aspartate transaminse level, IU/L	LE / AST	51.0	(33.0, 69.0)	51.0	(33.5, 69.0)	0	(0.5, 0)	52.0	(37.0, 87.0)	52.0	(37.0, 87.0)	0	(0, 0)	0.06	0.06
Albumin level	LE / ALB	4.0	(3.7, 4.2)	4.1	(3.8, 4.3)	0.1	(0.1, 0.1)	4.2	(4.0, 4.4)	4.2	(4.0, 4.4)	0	(0, 0)	<0.0001	<0.0001
Total bilirubin level, mg/dL	LE / TBILI	0.6	(0.4, 0.8)	0.6	(0.4, 0.8)	0	(0, 0)	0.7	(0.5, 0.9)	0.7	(0.5, 0.9)	0	(0, 0)	0.0008	0.0007
Hemoglobin level, g/dL	LE / HGB	14.3	(13.5, 15.1)	14.3	(13.5, 15.1)	0	(0, 0)	15.1	(13.9, 15.9)	15.0	(13.9, 15.9)	-0.1	(0, 0)	<0.0001	<0.0001
White blood cell count, 10^3 cells/mm^3	LE / WBC	5.8	(4.7, 7.4)	5.8	(4.7, 7.5)	0	(0, 0.1)	6.2	(4.9, 7.4)	6.2	(4.9, 7.4)	0	(0, 0)	0.09	0.08
HCV RNA level, x10^6 IU/mL **	VLOAD / vload_itt	6.4	(5.6, 6.7)	6.4	(5.6, 6.7)	0	(0, 0)	6.5	(5.7, 6.8)	6.5	(5.7, 6.8)	0	(0, 0)	0.09	0.08

NOTE: Continuous variables are represented by median (25th percentile, 75th percentile). The Wilcoxon rank-sum test was used to compare distributions for continuous variables, Pearson's ?^{2} test was used to compare percentages.

* archived data

** HCV RNA level reported as $\log_{10}$, per E-Mail communication with DCC, 8/13/2007

Primary Outcome. The study publication reported that the primary end point, sustained virologic response (SVR), was achieved in 28% of AA compared with 52% of CA patients (*P* < 0.0001). Similarly, analyses of archived data indicated that SVR was achieved in 27% of AA compared with 52% of CA patients (*P* < 0.0001). Selected univariable associations with SVR were analyzed for purposes of this DSIC. Results are presented in DSIC Table 2. Analyses of archived data indicated the following factors were associated with higher SVR: CA race, female sex, more than high school education, and lower weight. A history of diabetes was associated with lower response rates. Some laboratory measures were found to be possibly associated with lower response rates, including serum AST and WBC. Relative risks, 95% confidence intervals, and *P*-values for these selected univariable associations were very similar to those published by the study group (DSIC Table 2).

DSIC Table 2[†]. Relationship between SVR and variables of interest: Selected Univariable Analyses[7]

[†] *analyses selected from publication Table 3, p. 475.*

	Univariable Analysis						
	Archived		Published			P -value	
Variable	Relative Risk	(95% CI)	Relative Risk	(95% CI)	RR: Percent difference	Archived	Published
CA race	1.93	(1.48 - 2.51)	1.89	(1.46 - 2.46)	-2.1%	<0.0001	<0.0001
Male sex	0.74	(0.59 - 0.94)	0.73	(0.58 - 0.93)	-1.4%	0.02	0.01
Age, per 5 years	0.96	(0.89 - 1.03)	0.97	(0.89 - 1.03)	1.0%	0.23	0.24
Education (<high school)	0.71	(0.55 - 0.93)	0.72	(0.55 - 0.94)	1.4%	0.01	0.02
Weight (kg) per 5 kg	0.94	(0.91 - 0.97)	0.94	(0.91 - 0.97)	0.0%	0.0003	0.0004
BMI, kg/m^2	0.98	(0.96 - 0.9995)	0.98	(0.96 - 1.00)	0.0%	0.046	0.07
History of diabetes	0.49	(0.26 - 0.92)	0.49	(0.26 - 0.91)	0.0%	0.03	0.02
History of hypertension	0.73	(0.55 - 0.98)	0.76	(0.57 - 1.01)	3.9%	0.03	0.06
Current smoking	1.14	(0.89 - 1.45)	1.13	(0.89 - 1.45)	-0.9%	0.30	0.31
Alanine transaminase level (IU) per 100 IU	1.05	(0.91 - 1.22)	1.05	(0.90 - 1.22)	0.0%	0.49	0.53
Asparate transaminase level (IU) per 100 IU	0.75	(0.55 - 1.03)	0.75	(0.55 - 1.02)	0.0%	0.07	0.06
Hemoglobin, g/DL	0.95	(0.87 - 1.04)	0.95	(0.87 - 1.04)	0.0%	0.30	0.26
White blood cells, per 10^3 cells/mm^3	1.05	(0.993 - 1.10)	1.05	(0.996 - 1.11)	0.0%	0.09	0.07
Baseline viral level, $\log_{10}$ IU/mL	0.78	(0.71 - 0.86)	0.76	(0.66 - 0.88)	-2.6%	<0.0001	0.0002

[7] For purposes of the DSIC, generalized estimating equations (GEE) (*PC-SAS: Proc Genmod,* where *dist=binomial, link=log*) were used to estimate relative risk of univariable associations with SVR. The study publication does not explicitly state the statistical methods used to assess univariable associations. Small differences between archived and published results may be expected, due to differences in methods of estimation.

Notes

(1) Eight of 34 raw datasets were used for the current analysis. Additional datasets will be analyzed as part of dataset integrity checks for other main publications.

References

[1] Virahep-C Study Group (Hari S. Conjeevaram et al). Peginterferon and Ribavirin Treatment in African American and Caucasian American Patients With Hepatitis C Genotype 1. **Gastroenterology** 2006;131:470–477.

Attachment 1

“The full text of the article referenced will be provided to approved data requestors along with the data archived.”

Hari S. Conjeevaram, Michael W. Fried, Lennox J. Jeffers, Norah A. Terrault, Thelma E. Wiley-Lucas, Nezam Afdhal, Robert S. Brown, Steven H. Belle, Jay H. Hoofnagle, David E. Kleiner, and Charles D. Howell, for the Virahep-C Study Group. **Peginterferon and Ribavirin Treatment in African American and Caucasian American Patients With Hepatitis C Genotype 1**.
Gastroenterology 2006;131:470–477

Attachment 2

SAS 9.1 Log for programming code submitted for the replication of results in Table 1 “Baseline Characteristics of Virahep-C Participants” Conjeevaram et.al. 2006

```
1                                         The SAS System

NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) 9.1 (TS1M3)
      Licensed to RESEARCH TRIANGLE INSTITUTE, Site 0047670011.
NOTE: This session is executing on the XP_PRO  platform.



NOTE: SAS 9.1.3 Service Pack 3

NOTE: SAS initialization used:
      real time           2.30 seconds
      cpu time            0.32 seconds

1          options ps=55 ls=78 nonumber formchar='|----|+\---+=|-^<>*' mprint
orientation=portrait
1        ! ;
2
3          libname main 'Z:\05_Users\Sylvia\Virahep-C\main_study\mainstudy';
NOTE: Libref MAIN was successfully assigned as follows:
      Engine:        V9
      Physical Name: Z:\05_Users\Sylvia\Virahep-C\main_study\mainstudy
4          %include 'Z:\05_Users\Sylvia\Virahep-C\vhepfmt.sas';
NOTE: Format ACTION has been output.
NOTE: Format AEVENT has been output.
NOTE: Format ALC has been output.
NOTE: Format ALL5NO has been output.
NOTE: Format ALL6NO has been output.
NOTE: Format ALWNEV has been output.
NOTE: Format BDYPN has been output.
NOTE: Format BETTER has been output.
NOTE: Format BIN has been output.
NOTE: Format BXSC has been output.
NOTE: Format CESD has been output.
NOTE: Format COMPBY has been output.
NOTE: Format DCHRS has been output.
NOTE: Format DIAB has been output.
NOTE: Format DISCD has been output.
NOTE: Format DRUGDS has been output.
NOTE: Format DRUGRL has been output.
NOTE: Format EDUC has been output.
NOTE: Format EVTYPE has been output.
NOTE: Format EXCPOR has been output.
NOTE: Format GFBC has been output.
NOTE: Format GRF has been output.
NOTE: Format GI has been output.
NOTE: Format HCVRES has been output.
NOTE: Format HERBCAT has been output.
NOTE: Format HISP has been output.
NOTE: Format HL1YR has been output.
NOTE: Format INCDEC has been output.
NOTE: Format INS has been output.
NOTE: Format LIMIT has been output.
NOTE: Format MARST has been output.
NOTE: Format MEDDC has been output.
```

```
NOTE: Format MISINT has been output.
NOTE: Format MISSM has been output.
NOTE: Format MISRIB has been output.
NOTE: Format NEVALL has been output.
NOTE: Format NEVER has been output.
NOTE: Format NEVOFT has been output.
NOTE: Format NEVWK has been output.
NOTE: Format NONALL has been output.
NOTE: Format NOTEXT has been output.
NOTE: Format POSNEG has been output.
NOTE: Format REMEM has been output.
NOTE: Format RESLV has been output.
NOTE: Format RISK has been output.
NOTE: Format SATDIS has been output.
NOTE: Format SATISF has been output.
NOTE: Format SEVER has been output.
NOTE: Format SEX has been output.
NOTE: Format SEXFUNC has been output.
NOTE: Format SMOKE has been output.
NOTE: Format SOC has been output.
NOTE: Format SOURCE has been output.
NOTE: Format STATIN has been output.
NOTE: Format STDYDC has been output.
NOTE: Format TIMEOP has been output.
NOTE: Format TMDC has been output.
NOTE: Format TMPT has been output.
NOTE: Format TMPTT has been output.
NOTE: Format TMPTDW has been output.
NOTE: Format TRFAL has been output.
NOTE: Format TTMPT has been output.
NOTE: Format USA has been output.
NOTE: Format VISIT has been output.
NOTE: Format VQUALDET has been output.
NOTE: Format WORK has been output.
NOTE: Format YESNO has been output.

NOTE: PROCEDURE FORMAT used (Total process time):
      real time           0.16 seconds
      cpu time            0.03 seconds


830        libname analysis 'Z:\05_Users\Sylvia\Virahep-C\main_study\analysis';
NOTE: Libref ANALYSIS was successfully assigned as follows:
      Engine:        V9
      Physical Name: Z:\05_Users\Sylvia\Virahep-C\main_study\analysis
831
832        * define sample of 401 and perform bsln analyses *;
833        data sc; set main.sc;

NOTE: There were 546 observations read from the data set MAIN.SC.
NOTE: The data set WORK.SC has 546 observations and 52 variables.
NOTE: DATA statement used (Total process time):
      real time           0.50 seconds
```

```
      cpu time            0.02 seconds


834          proc sort; by vhcid rsc;

NOTE: There were 546 observations read from the data set WORK.SC.
NOTE: The data set WORK.SC has 546 observations and 52 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


835          data sc; set sc; by vhcid rsc; if last.vhcid;
836          keep vhcid rsc cons; *cons = gave informed consent *;
837

NOTE: There were 546 observations read from the data set WORK.SC.
NOTE: The data set WORK.SC has 536 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


838        data demo; set main.se; keep vhcid rsc male racew--raceos wgtkg mxdiab mxlla
mxbp dob
838      ! doe age hgtin hgtcm bmi;
839          wgtkg=wgtlb/2.2;
840          if sex=1 then male=1;
841          else if sex=2 then male=0;
842          dob=mdy(dobm,dobd,doby);
843          doe=mdy(doem,doed,doey);
844          format dob doe MMDDYY10.;
845          age=round((doe-dob)/365.25,0.1);
846          hgtcm=round((hgtin*2.54),0.1);
847          bmi=wgtkg/(hgtcm*hgtcm/10000);  ***weight kg / ht2 in m***;

NOTE: Missing values were generated as a result of performing an operation on missing
values.
      Each place is given by: (Number of times) at (Line):(Column).
      17 at 839:14   30 at 846:9    30 at 846:21   5 at 847:12    30 at 847:19
NOTE: There were 552 observations read from the data set MAIN.SE.
NOTE: The data set WORK.DEMO has 552 observations and 20 variables.
NOTE: DATA statement used (Total process time):
      real time           1.34 seconds
      cpu time            0.02 seconds


848        proc sort; by vhcid rsc; * rsc = rescreen *;

NOTE: There were 552 observations read from the data set WORK.DEMO.
NOTE: The data set WORK.DEMO has 552 observations and 20 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds
```

```
849        data demo; set demo; by vhcid rsc;
850          if first.vhcid;
851

NOTE: There were 552 observations read from the data set WORK.DEMO.
NOTE: The data set WORK.DEMO has 536 observations and 20 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


852        data screen; set main.sq; keep vhcid educ evsmk cursmk alc alcbe alcwi alclq;

NOTE: There were 430 observations read from the data set MAIN.SQ.
NOTE: The data set WORK.SCREEN has 430 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.40 seconds
      cpu time            0.02 seconds


853        proc sort; by vhcid;
854

NOTE: There were 430 observations read from the data set WORK.SCREEN.
NOTE: The data set WORK.SCREEN has 430 observations and 8 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds


855        data labbase; set main.le; if tmpt=3;

NOTE: There were 7345 observations read from the data set MAIN.LE.
NOTE: The data set WORK.LABBASE has 401 observations and 80 variables.
NOTE: DATA statement used (Total process time):
      real time           6.91 seconds
      cpu time            0.06 seconds


856          proc sort; by vhcid;

NOTE: There were 401 observations read from the data set WORK.LABBASE.
NOTE: The data set WORK.LABBASE has 401 observations and 80 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


857        data hcv_base; set main.vload;
858           if tmpt=3;

NOTE: There were 5247 observations read from the data set MAIN.VLOAD.
```

```
NOTE: The data set WORK.HCV_BASE has 401 observations and 13 variables.
NOTE: DATA statement used (Total process time):
      real time           1.11 seconds
      cpu time            0.00 seconds



859          proc sort; by vhcid;
860
861        * hepatic steatosis / fibrosis measures *;

NOTE: There were 401 observations read from the data set WORK.HCV_BASE.
NOTE: The data set WORK.HCV_BASE has 401 observations and 13 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds



862        data hepa; set main.pe;

NOTE: There were 429 observations read from the data set MAIN.PE.
NOTE: The data set WORK.HEPA has 429 observations and 50 variables.
NOTE: DATA statement used (Total process time):
      real time           0.43 seconds
      cpu time            0.01 seconds



863          proc sort; by vhcid;
864

NOTE: There were 429 observations read from the data set WORK.HEPA.
NOTE: The data set WORK.HEPA has 429 observations and 50 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds



865        data demobase; merge sc demo screen labbase(in=in1) hcv_base(in=in2) hepa;
866          by vhcid; if in1 or in2; * anyone with bsln lab or viral load data, n=403 *;
867          if vhcid^=295076; * found to be ineligible during course of study, even
though has bsln and f/up data (DCC email 8/13/07) *;
868          if cons=1; * 292043 withdrew consent at bsln(cons=0), has bsln vload data,
869                                  but has no bsln lab or followup data ;
870          * THIS GIVES US FINAL N OF 401 ELIGIBLE (PRESUMABLY -- SINCE HAVE BSLN DATA
871                                        AND CONSENTED TO STUDY) *;
872          if educ>4 then hsorless=0; * high school or less *;
873          else if .<educ<=4 then hsorless=1;
874          log10vload=log10(vload_itt);
875          currentsmk=cursmk;
876          if evsmk in (0,.B) then currentsmk=0; *cursmk was set to n/a when evsmk=0 *;
877
878          run;
NOTE: Missing values were generated as a result of performing an operation on missing
values.
      Each place is given by: (Number of times) at (Line):(Column).
```

```
      2 at 874:14
NOTE: There were 536 observations read from the data set WORK.SC.
NOTE: There were 536 observations read from the data set WORK.DEMO.
NOTE: There were 430 observations read from the data set WORK.SCREEN.
NOTE: There were 401 observations read from the data set WORK.LABBASE.
NOTE: There were 401 observations read from the data set WORK.HCV_BASE.
NOTE: There were 429 observations read from the data set WORK.HEPA.
NOTE: The data set WORK.DEMOBASE has 401 observations and 167 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


879
880          DATA ANALYSIS.DEMOBASE; SET DEMOBASE; RUN;

NOTE: There were 401 observations read from the data set WORK.DEMOBASE.
NOTE: The data set ANALYSIS.DEMOBASE has 401 observations and 167 variables.
NOTE: DATA statement used (Total process time):
      real time           1.51 seconds
      cpu time            0.04 seconds


881
882          *************************************************************;
883          * for table 1 (analysis of baseline factors) *;
884          * use continuity adjusted chisq test for proportions *;
885          * wilcoxon 2-group z-test for medians *;
886          *************************************************************;
887        proc printto print='Z:\05_Users\Sylvia\Virahep-C\main_study\vhepc_demo.out'
new;
888          title table 1: analysis of baseline factors;

NOTE: PROCEDURE PRINTTO used (Total process time):
      real time           0.14 seconds
      cpu time            0.00 seconds


889          proc freq data=demobase; tables racew*raceb/missing; run;

NOTE: There were 401 observations read from the data set WORK.DEMOBASE.
NOTE: The PROCEDURE FREQ printed page 1.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.07 seconds
      cpu time            0.01 seconds


890          proc freq data=demobase; tables (male hsorless mxdiab mxbp
currentsmk)*racew/chisq
890      ! exact; run;

NOTE: There were 401 observations read from the data set WORK.DEMOBASE.
NOTE: The PROCEDURE FREQ printed pages 2-7.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.05 seconds
```

```
      cpu time            0.03 seconds


891          proc univariate data=demobase plot normal; class racew; var age wgtkg bmi
alt ast alb
891      !  tbili hgb wbc log10vload; run;

NOTE: The PROCEDURE UNIVARIATE printed pages 8-77.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.28 seconds
      cpu time            0.05 seconds


892          proc npar1way wilcoxon data=demobase; class racew; var age wgtkg bmi alt ast
alb
892      ! tbili hgb wbc log10vload; run;

NOTE: There were 401 observations read from the data set WORK.DEMOBASE.
NOTE: The PROCEDURE NPAR1WAY printed pages 78-87.
NOTE: PROCEDURE NPAR1WAY used (Total process time):
      real time           0.04 seconds
      cpu time            0.03 seconds


893        proc printto; run;

NOTE: PROCEDURE PRINTTO used (Total process time):
      real time           0.08 seconds
      cpu time            0.00 seconds


894          ************************************************************;

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
NOTE: The SAS System used:
      real time           16.70 seconds
      cpu time            0.82 seconds
```

Attachment 3

SAS 9.1 Log for programming code submitted for the replication of selected results in Table 3 “Relationship Between SVR and Each Variable of Interest in Univariable and Multivariable Analysis” Conjeevaram et.al. 2006

```
NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) 9.1 (TS1M3)
      Licensed to RESEARCH TRIANGLE INSTITUTE, Site 0047670011.
NOTE: This session is executing on the XP_PRO  platform.



NOTE: SAS 9.1.3 Service Pack 3

NOTE: SAS initialization used:
      real time           3.93 seconds
      cpu time            0.32 seconds

1          options ps=55 ls=78 nonumber formchar='|----|+\---+=|-^<>*' mprint
orientation=portrait
1        ! ;
2
3          libname main 'Z:\05_Users\Sylvia\Virahep-C\main_study\mainstudy';
NOTE: Libref MAIN was successfully assigned as follows:
      Engine:        V9
      Physical Name: Z:\05_Users\Sylvia\Virahep-C\main_study\mainstudy
4          libname library 'Z:\05_Users\Sylvia\Virahep-C\main_study';
NOTE: Libref LIBRARY was successfully assigned as follows:
      Engine:        V9
      Physical Name: Z:\05_Users\Sylvia\Virahep-C\main_study
5          libname analysis 'Z:\05_Users\Sylvia\Virahep-C\main_study\analysis';
NOTE: Libref ANALYSIS was successfully assigned as follows:
      Engine:        V9
      Physical Name: Z:\05_Users\Sylvia\Virahep-C\main_study\analysis
6
7          DATA DEMOBASE; SET ANALYSIS.DEMOBASE; RUN;

NOTE: There were 401 observations read from the data set ANALYSIS.DEMOBASE.
NOTE: The data set WORK.DEMOBASE has 401 observations and 167 variables.
NOTE: DATA statement used (Total process time):
      real time           4.76 seconds
      cpu time            0.06 seconds



8
9            * construct SVR outcome *;
10         data hcv_fuwk24; set main.vload;
11            if tmpt=25; *f-up wk 24 *;

NOTE: There were 5247 observations read from the data set MAIN.VLOAD.
NOTE: The data set WORK.HCV_FUWK24 has 326 observations and 13 variables.
NOTE: DATA statement used (Total process time):
      real time           3.34 seconds
      cpu time            0.03 seconds


12         proc sort; by vhcid;
13          *note: DATASET VLOAD contains both qualitative and quantitative results *;
14
```

```
NOTE: There were 326 observations read from the data set WORK.HCV_FUWK24.
NOTE: The data set WORK.HCV_FUWK24 has 326 observations and 13 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


15         data all24_fu; merge demobase(in=in1 keep=vhcid age bmi male racew raceb
16                                  hsorless mxdiab mxbp currentsmk
17                                  wgtkg wbc alt ast hgb log10vload vload vload_itt
18                                  gluf insf fatsc2
19                                  rename=(log10vload=blog10vload vload=bvload
20                                          vload_itt=bvload_itt))
21                             hcv_fuwk24(in=in2);
22             by vhcid;
23          * Defn of primary endpoint:  an SVR defined as the absence of detectable
24                 HCV RNA in serum 24 weeks after therapy was stopped... At week 24 and
25                 time points thereafter, missing HCV RNA data were considered a
26                 nonresponse (Conjeevaram 2006, p. 471);
27
28          *initially set SVR to VQUAL=0 at f-up wk 24 *;
29             svr=1-vqual;
30
31           if in1 and not in2 then do; svr=0; flag1=1; end;
32              *no data at fuwk 24 -- set to non-response *;
33           if in2 and vqual in (.,.C) then do;
34             if vload_itt>599 then svr=0; *detectable viral load
35                                              using quantitative tests *;
36             else do; svr=0; flag2=1; end;
37          * vload_itt=599 is an unknown response, so set to non-response*;
38           end;
39
40           if in1; * keep all eligible in dataset *;
41           if racew=1 then race='W'; else if racew=0 then race='B';
42           age_5=age/5;
43           wgtkg_5=wgtkg/5;
44           alt_100=alt/100;
45           ast_100=ast/100;
46           * steatosis outcome (based on FATSC2, in main.pe ) *;
47           steat=fatsc2;
48           if fatsc2 in (2,3,4) then steat=2;
49           homa=(insf*gluf/18)/22.5;
50
51           * per email DCC (8/13/07): if HCV data not available at fuwk24,
52                                       then glean from later visits *;

NOTE: Missing values were generated as a result of performing an operation on missing
values.
      Each place is given by: (Number of times) at (Line):(Column).
      186 at 29:10   3 at 43:16     2 at 44:14     2 at 45:14     59 at 49:13
NOTE: There were 401 observations read from the data set WORK.DEMOBASE.
NOTE: There were 326 observations read from the data set WORK.HCV_FUWK24.
NOTE: The data set WORK.ALL24_FU has 401 observations and 43 variables.
```

```
NOTE: DATA statement used (Total process time):
      real time           0.95 seconds
      cpu time            0.01 seconds


53         data postfu_wk24; set main.vload;
54             if tmpt in (26,27,33);
55

NOTE: There were 5247 observations read from the data set MAIN.VLOAD.
NOTE: The data set WORK.POSTFU_WK24 has 220 observations and 13 variables.
NOTE: DATA statement used (Total process time):
      real time           3.37 seconds
      cpu time            0.03 seconds


55       !  proc sort; by vhcid tmpt;
56

NOTE: There were 220 observations read from the data set WORK.POSTFU_WK24.
NOTE: The data set WORK.POSTFU_WK24 has 220 observations and 13 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


57           proc freq data=all24_fu; where flag1=1; tables vhcid/out=nohcv_fu24;

NOTE: There were 75 observations read from the data set WORK.ALL24_FU.
      WHERE flag1=1;
NOTE: The data set WORK.NOHCV_FU24 has 75 observations and 3 variables.
NOTE: The PROCEDURE FREQ printed pages 1-2.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.12 seconds
      cpu time            0.03 seconds


58           proc freq data=all24_fu; where flag2=1; tables vhcid/out=misshcv_fu24;
59           run;

NOTE: There were 3 observations read from the data set WORK.ALL24_FU.
      WHERE flag2=1;
NOTE: The data set WORK.MISSHCV_FU24 has 3 observations and 3 variables.
NOTE: The PROCEDURE FREQ printed page 3.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


60             * pick up later data for those with no data at fuwk24 *;
61           data nohcv_fu24_post; merge nohcv_fu24(in=in1) postfu_wk24(in=in2);
62             by vhcid; if in1 and in2;
63             svr=1-vqual;
```

```
64             if vqual=. then do; if vload_itt>599 then svr=0; end; run;

NOTE: Missing values were generated as a result of performing an operation on missing
values.
      Each place is given by: (Number of times) at (Line):(Column).
      2 at 63:10
NOTE: There were 75 observations read from the data set WORK.NOHCV_FU24.
NOTE: There were 220 observations read from the data set WORK.POSTFU_WK24.
NOTE: The data set WORK.NOHCV_FU24_POST has 5 observations and 16 variables.
NOTE: DATA statement used (Total process time):
      real time           0.67 seconds
      cpu time            0.01 seconds


65          /* * check results: ;
66          proc print; by vhcid; id vhcid; var tmpt vqual--vload_itt svr;
67               run; */
68             * pick up later data for those missing data at fuwk24 *;
69           data misshcv_fu24_post; merge misshcv_fu24(in=in1) postfu_wk24(in=in2);
70             by vhcid; if in1 and in2;
71             svr=1-vqual;
72             if vqual=. then do; if vload_itt>599 then svr=0; end; run;

NOTE: There were 3 observations read from the data set WORK.MISSHCV_FU24.
NOTE: There were 220 observations read from the data set WORK.POSTFU_WK24.
NOTE: The data set WORK.MISSHCV_FU24_POST has 3 observations and 16 variables.
NOTE: DATA statement used (Total process time):
      real time           0.64 seconds
      cpu time            0.00 seconds


73          /* * check results: ;
74          proc print data=misshcv_fu24_post; by vhcid;
75               id vhcid; var tmpt vqual--vload_itt svr; run; */
76
77         data all24_fu2; merge all24_fu nohcv_fu24_post(keep=vhcid svr)
78                        misshcv_fu24_post(keep=vhcid svr);
79           by vhcid;
80           run;

NOTE: There were 401 observations read from the data set WORK.ALL24_FU.
NOTE: There were 5 observations read from the data set WORK.NOHCV_FU24_POST.
NOTE: There were 3 observations read from the data set WORK.MISSHCV_FU24_POST.
NOTE: The data set WORK.ALL24_FU2 has 401 observations and 43 variables.
NOTE: DATA statement used (Total process time):
      real time           0.63 seconds
      cpu time            0.00 seconds


81
82           * final analysis of race by SVR *;
83         proc freq data=all24_fu2; tables race*svr/chisq exact; run;

NOTE: There were 401 observations read from the data set WORK.ALL24_FU2.
```

```
NOTE: The PROCEDURE FREQ printed page 4.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds



84
85            ** use log-binomial for all continuous to get RR **;
86            * for replication of table 2 *;
87          %macro relrisks(covar);
88            proc genmod data=all24_fu2 descending;
89            model svr=&covar/dist = binomial link=log type3;
90            estimate "Beta &covar" &covar 1 -1/ exp;
91            run;
92          %mend;
93
94          * the warning " more coefficients specified than levels" pops up when an
95            estimate of the beta coefficient (RR) -- with contrast levels 1 -1 --
96            is requested for a trend test, this does not seem to affect estimates *;
97          proc printto
98            print='Z:\05_Users\Sylvia\Virahep-C\main_study\vhepc_svr.out' new;
99          title replication of selected results in table 3: SVR<=>variables of interest;
100         %relrisks(racew);

NOTE: PROCEDURE PRINTTO used (Total process time):
      real time           0.15 seconds
      cpu time            0.00 seconds


MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=racew/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta racew" racew 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect RACEW.  Some coefficients
will be
         ignored.
NOTE: The PROCEDURE GENMOD printed pages 5-6.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


101         %relrisks(male);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=male/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta male" male 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
```

```
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect male.  Some coefficients will
be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 7-8.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


102        %relrisks(age_5);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=age_5/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta age_5" age_5 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect age_5.  Some coefficients
will be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 9-10.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.03 seconds
      cpu time            0.01 seconds


103        %relrisks(hsorless);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=hsorless/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta hsorless" hsorless 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect hsorless.  Some coefficients
will be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 11-12.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


104        %relrisks(wgtkg_5);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=wgtkg_5/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta wgtkg_5" wgtkg_5 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect wgtkg_5.  Some coefficients
will be
```

```
         ignored.
NOTE: The PROCEDURE GENMOD printed pages 13-14.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


105        %relrisks(bmi);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=bmi/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta bmi" bmi 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect bmi.  Some coefficients will
be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 15-16.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


106        %relrisks(mxdiab);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=mxdiab/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta mxdiab" mxdiab 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect MXDIAB.  Some coefficients
will be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 17-18.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


107        %relrisks(mxbp);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=mxbp/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta mxbp" mxbp 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect MXBP.  Some coefficients will
be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 19-20.
```

```
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.03 seconds
      cpu time            0.01 seconds


108        %relrisks(currentsmk);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=currentsmk/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta currentsmk" currentsmk 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect currentsmk.  Some
coefficients will          be ignored.
NOTE: The PROCEDURE GENMOD printed pages 21-22.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


109        %relrisks(alt_100);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=alt_100/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta alt_100" alt_100 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect alt_100.  Some coefficients
will be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 23-24.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.03 seconds
      cpu time            0.01 seconds


110        %relrisks(ast_100);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=ast_100/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta ast_100" ast_100 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect ast_100.  Some coefficients
will be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 25-26.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
```

```
      cpu time            0.01 seconds


111        %relrisks(hgb);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=hgb/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta hgb" hgb 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect HGB.  Some coefficients will
be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 27-28.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


112        %relrisks(wbc);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=wbc/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta wbc" wbc 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect WBC.  Some coefficients will
be          ignored.
NOTE: The PROCEDURE GENMOD printed pages 29-30.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


113        %relrisks(blog10vload);
MPRINT(RELRISKS):   proc genmod data=all24_fu2 descending;
MPRINT(RELRISKS):   model svr=blog10vload/dist = binomial link=log type3;
MPRINT(RELRISKS):   estimate "Beta blog10vload" blog10vload 1 -1/ exp;
MPRINT(RELRISKS):   run;

NOTE: PROC GENMOD is modeling the probability that svr='1'.
NOTE: Algorithm converged.
NOTE: The scale parameter was held fixed.
WARNING: More coefficients than levels specified for effect blog10vload.  Some
coefficients will          be ignored.
NOTE: The PROCEDURE GENMOD printed pages 31-32.
NOTE: PROCEDURE GENMOD used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds
```

```
114        proc printto; run;

NOTE: PROCEDURE PRINTTO used (Total process time):
      real time           0.07 seconds
      cpu time            0.00 seconds


NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
NOTE: The SAS System used:
      real time           21.46 seconds
      cpu time            0.93 seconds
```

Attachment 4

SAS 9.1 Output for programming code submitted for the replication of results in Table 1 “Baseline Characteristics of Virahep-C Participants” Conjeevaram et.al. 2006

```
                       table 1: analysis of baseline factors

                                 The FREQ Procedure

                              Table of RACEW by RACEB

                        RACEW(White or Caucasian)
                                  RACEB(Black or African American)
                        Frequency|
                        Percent  |
                        Row Pct  |
                        Col Pct  |No      |Yes     |  Total
                        ---------+--------+--------+
                        No       |      0 |    196 |    196
                                 |   0.00 |  48.88 |  48.88
                                 |   0.00 | 100.00 |
                                 |   0.00 | 100.00 |
                        ---------+--------+--------+
                        Yes      |    205 |      0 |    205
                                 |  51.12 |   0.00 |  51.12
                                 | 100.00 |   0.00 |
                                 | 100.00 |   0.00 |
                        ---------+--------+--------+
                        Total         205      196      401
                                    51.12    48.88   100.00
```

```
                 table 1: analysis of baseline factors


                          The FREQ Procedure

                        Table of male by RACEW

                  male      RACEW(White or Caucasian)

                  Frequency|
                  Percent  |
                  Row Pct  |
                  Col Pct  |No      |Yes     |  Total
                  ---------+--------+--------+
                         0 |     69 |     71 |    140
                           |  17.21 |  17.71 |  34.91
                           |  49.29 |  50.71 |
                           |  35.20 |  34.63 |
                  ---------+--------+--------+
                         1 |    127 |    134 |    261
                           |  31.67 |  33.42 |  65.09
                           |  48.66 |  51.34 |
                           |  64.80 |  65.37 |
                  ---------+--------+--------+
                  Total         196      205      401
                              48.88    51.12   100.00


                 Statistics for Table of male by RACEW

         Statistic                     DF       Value      Prob
         ------------------------------------------------------
         Chi-Square                     1      0.0143    0.9047
         Likelihood Ratio Chi-Square    1      0.0143    0.9047
         Continuity Adj. Chi-Square     1      0.0002    0.9881
         Mantel-Haenszel Chi-Square     1      0.0143    0.9049
         Phi Coefficient                       0.0060
         Contingency Coefficient               0.0060
         Cramer's V                            0.0060


                          Fisher's Exact Test
                   ----------------------------------
                   Cell (1,1) Frequency (F)        69
                   Left-sided Pr <= F          0.5889
                   Right-sided Pr >= F         0.4940

                   Table Probability (P)       0.0828
                   Two-sided Pr <= P           0.9169

                           Sample Size = 401
```

```
              table 1: analysis of baseline factors

                        The FREQ Procedure

                    Table of hsorless by RACEW

                hsorless     RACEW(White or Caucasian)

                Frequency|
                Percent  |
                Row Pct  |
                Col Pct  |No      |Yes     |  Total
                ---------+--------+--------+
                       0 |    104 |    132 |    236
                         |  26.33 |  33.42 |  59.75
                         |  44.07 |  55.93 |
                         |  54.17 |  65.02 |
                ---------+--------+--------+
                       1 |     88 |     71 |    159
                         |  22.28 |  17.97 |  40.25
                         |  55.35 |  44.65 |
                         |  45.83 |  34.98 |
                ---------+--------+--------+
                Total         192      203      395
                            48.61    51.39   100.00

                        Frequency Missing = 6
```

```
                    table 1: analysis of baseline factors


                             The FREQ Procedure

                  Statistics for Table of hsorless by RACEW

           Statistic                     DF       Value      Prob
           ------------------------------------------------------
           Chi-Square                     1      4.8371    0.0279
           Likelihood Ratio Chi-Square    1      4.8446    0.0277
           Continuity Adj. Chi-Square     1      4.3961    0.0360
           Mantel-Haenszel Chi-Square     1      4.8248    0.0281
           Phi Coefficient                      -0.1107
           Contingency Coefficient               0.1100
           Cramer's V                           -0.1107


                           Fisher's Exact Test
                     ----------------------------------
                     Cell (1,1) Frequency (F)       104
                     Left-sided Pr <= F          0.0180
                     Right-sided Pr >= F         0.9894

                     Table Probability (P)       0.0073
                     Two-sided Pr <= P           0.0313

                        Effective Sample Size = 395
                           Frequency Missing = 6


                         Table of MXDIAB by RACEW

                    MXDIAB(Med history diabetes)
                              RACEW(White or Caucasian)
                    Frequency|
                    Percent  |
                    Row Pct  |
                    Col Pct  |No      |Yes     |  Total
                    ---------+--------+--------+
                    No       |    166 |    196 |    362
                             |  41.40 |  48.88 |  90.27
                             |  45.86 |  54.14 |
                             |  84.69 |  95.61 |
                    ---------+--------+--------+
                    Yes      |     30 |      9 |     39
                             |   7.48 |   2.24 |   9.73
                             |  76.92 |  23.08 |
                             |  15.31 |   4.39 |
                    ---------+--------+--------+
                    Total         196      205      401
                                48.88    51.12   100.00
```

```
        table 1: analysis of baseline factors

                  The FREQ Procedure

       Statistics for Table of MXDIAB by RACEW

Statistic                     DF       Value      Prob
------------------------------------------------------
Chi-Square                     1     13.5987    0.0002
Likelihood Ratio Chi-Square    1     14.2166    0.0002
Continuity Adj. Chi-Square     1     12.3839    0.0004
Mantel-Haenszel Chi-Square     1     13.5648    0.0002
Phi Coefficient                      -0.1842
Contingency Coefficient               0.1811
Cramer's V                           -0.1842


                 Fisher's Exact Test
          -----------------------------------
          Cell (1,1) Frequency (F)        166
          Left-sided Pr <= F        1.694E-04
          Right-sided Pr >= F          1.0000

          Table Probability (P)     1.297E-04
          Two-sided Pr <= P         2.969E-04

                  Sample Size = 401


                Table of MXBP by RACEW

         MXBP(Med hx hypertension)
                   RACEW(White or Caucasian)
         Frequency|
         Percent  |
         Row Pct  |
         Col Pct  |No      |Yes     |  Total
         ---------+--------+--------+
         No       |    111 |    162 |    273
                  |  27.68 |  40.40 |  68.08
                  |  40.66 |  59.34 |
                  |  56.63 |  79.02 |
         ---------+--------+--------+
         Yes      |     85 |     43 |    128
                  |  21.20 |  10.72 |  31.92
                  |  66.41 |  33.59 |
                  |  43.37 |  20.98 |
         ---------+--------+--------+
         Total         196      205      401
                     48.88    51.12   100.00
```

```
                  table 1: analysis of baseline factors


                            The FREQ Procedure

                  Statistics for Table of MXBP by RACEW

          Statistic                     DF       Value      Prob
          ------------------------------------------------------
          Chi-Square                     1     23.1184    <.0001
          Likelihood Ratio Chi-Square    1     23.4215    <.0001
          Continuity Adj. Chi-Square     1     22.0995    <.0001
          Mantel-Haenszel Chi-Square     1     23.0607    <.0001
          Phi Coefficient                      -0.2401
          Contingency Coefficient               0.2335
          Cramer's V                           -0.2401


                           Fisher's Exact Test
                    ----------------------------------
                    Cell (1,1) Frequency (F)       111
                    Left-sided Pr <= F       1.129E-06
                    Right-sided Pr >= F         1.0000

                    Table Probability (P)    7.536E-07
                    Two-sided Pr <= P        2.050E-06

                            Sample Size = 401


                       Table of currentsmk by RACEW

                   currentsmk     RACEW(White or Caucasian)

                   Frequency|
                   Percent  |
                   Row Pct  |
                   Col Pct  |No      |Yes     |  Total
                   ---------+--------+--------+
                          0 |    112 |    133 |    245
                            |  28.50 |  33.84 |  62.34
                            |  45.71 |  54.29 |
                            |  58.95 |  65.52 |
                   ---------+--------+--------+
                          1 |     78 |     70 |    148
                            |  19.85 |  17.81 |  37.66
                            |  52.70 |  47.30 |
                            |  41.05 |  34.48 |
                   ---------+--------+--------+
                   Total         190      203      393
                               48.35    51.65   100.00

                           Frequency Missing = 8
```

```
                    The FREQ Procedure

        Statistics for Table of currentsmk by RACEW

 Statistic                     DF       Value      Prob
 ------------------------------------------------------
 Chi-Square                     1      1.8044    0.1792
 Likelihood Ratio Chi-Square    1      1.8047    0.1791
 Continuity Adj. Chi-Square     1      1.5354    0.2153
 Mantel-Haenszel Chi-Square     1      1.7998    0.1797
 Phi Coefficient                      -0.0678
 Contingency Coefficient               0.0676
 Cramer's V                           -0.0678


                   Fisher's Exact Test
           ----------------------------------
           Cell (1,1) Frequency (F)       112
           Left-sided Pr <= F          0.1076
           Right-sided Pr >= F         0.9261

           Table Probability (P)       0.0338
           Two-sided Pr <= P           0.2113

               Effective Sample Size = 393
                  Frequency Missing = 8
```

```
                   table 1: analysis of baseline factors

                          The UNIVARIATE Procedure
                             Variable:  age
                               RACEW = No

                                 Moments

N                         196    Sum Weights                196
Mean               49.0673469    Sum Observations        9617.2
Std Deviation      7.19345537    Variance            51.7458001
Skewness           0.19652272    Kurtosis            0.98649779
Uncorrected SS      481980.92    Corrected SS         10090.431
Coeff Variation    14.6603715    Std Error Mean      0.51381824


                        Basic Statistical Measures

         Location                    Variability

     Mean     49.06735     Std Deviation            7.19346
     Median   49.15000     Variance                51.74580
     Mode     49.20000     Range                   42.80000
                           Interquartile Range      7.70000


                        Tests for Location: Mu0=0

        Test           -Statistic-    -----p Value------

        Student's t    t  95.49553    Pr > |t|    <.0001
        Sign           M        98    Pr >= |M|   <.0001
        Signed Rank    S      9653    Pr >= |S|   <.0001


                         Quantiles (Definition 5)

                          Quantile      Estimate

                         10 0% Max        70.40
                          99%             69.50
                          95%             61.70
                          90%             57.90
                          75% Q3          53.00
                          50% Median      49.15
                          25% Q1          45.30
                          10%             39.80
                          5%              35.90
                          1%              31.50
                          0% Min          27.60
```

```
                      table 1: analysis of baseline factors


                            The UNIVARIATE Procedure
                                 Variable:  age
                                   RACEW = No

                              Extreme Observations

                       ----Lowest----        ----Highest---

                       Value      Obs        Value      Obs

                        27.6       86         67.4      135
                        31.5      145         68.7      147
                        32.7      157         68.9      312
                        33.9       49         69.5       78
                        34.1      205         70.4      151
```

```
                  table 1: analysis of baseline factors


                         The UNIVARIATE Procedure
                             Variable:  age
                              RACEW = Yes

                                Moments

    N                         205    Sum Weights                205
    Mean               47.3253659    Sum Observations        9701.7
    Std Deviation      8.50517377    Variance            72.3379809
    Skewness           -0.3471691    Kurtosis            0.97855102
    Uncorrected SS      473893.45    Corrected SS        14756.9481
    Coeff Variation    17.9717021    Std Error Mean       0.5940271


                      Basic Statistical Measures

         Location                    Variability

     Mean     47.32537     Std Deviation            8.50517
     Median   48.20000     Variance                72.33798
     Mode     56.70000     Range                   48.20000
                           Interquartile Range      9.50000


                     Tests for Location: Mu0=0

          Test           -Statistic-    -----p Value------

          Student's t    t   79.6687    Pr > |t|    <.0001
          Sign           M     102.5    Pr >= |M|   <.0001
          Signed Rank    S   10557.5    Pr >= |S|   <.0001


                       Quantiles (Definition 5)

                        Quantile      Estimate

                        100% Max          70.7
                        99%               69.6
                        95%               59.2
                        90%               56.7
                        75% Q3            52.4
                        50% Median        48.2
                        25% Q1            42.9
                        10%               37.3
                        5%                31.8
                        1%                24.1
                        0% Min            22.5
```

```
                      table 1: analysis of baseline factors


                            The UNIVARIATE Procedure
                                 Variable:  age
                                  RACEW = Yes

                              Extreme Observations

                       ----Lowest----        ----Highest---

                       Value      Obs        Value      Obs

                        22.5      369         64.3      111
                        23.1      367         68.6      196
                        24.1      164         69.6      113
                        24.4      326         70.4      234
                        25.4       56         70.7       21
```

```
             table 1: analysis of baseline factors


                    The UNIVARIATE Procedure
                        Variable:  wgtkg
                           RACEW = No

                            Moments

N                              193    Sum Weights                193
Mean                    91.0174282    Sum Observations    17566.3636
Std Deviation            19.628771    Variance             385.28865
Skewness                0.83374019    Kurtosis            0.84820102
Uncorrected SS          1672820.66    Corrected SS        73975.4207
Coeff Variation         21.5659477    Std Error Mean      1.41290985


                   Basic Statistical Measures

         Location                    Variability

     Mean     91.01743     Std Deviation            19.62877
     Median   87.27273     Variance                385.28865
     Mode     74.54545     Range                   102.27273
                           Interquartile Range      22.72727


                   Tests for Location: Mu0=0

        Test           -Statistic-    -----p Value------

        Student's t    t  64.41843    Pr > |t|    <.0001
        Sign           M      96.5    Pr >= |M|   <.0001
       Si gned Rank    S    9360.5    Pr >= |S|   <.0001


                    Quantiles (Definition 5)

                    Quantile       Estimate

                    100% Max       157.2727
                    99%            156.8182
                    95%            133.6364
                    90%            114.5455
                    75% Q3         100.9091
                    50% Median      87.2727
                    25% Q1          78.1818
                    10%             70.0000
                    5%              62.7273
                   1%               55.9091
                    0% Min          55.0000
```

```
                    table 1: analysis of baseline factors

                         The UNIVARIATE Procedure
                            Variable:  wgtkg
                              RACEW = No

                          Extreme Observations

                   ------Lowest-----        -----Highest-----

                      Value      Obs            Value      Obs

                    55.0000       49          138.182      341
                    55.9091      392          139.091      368
                    55.9091      315          148.182       82
                    57.2727       33          156.818       18
                    59.0909      172          157.273      202


                             Missing Values

                                      -----Percent Of-----
                  Missing                            Missing
                    Value       Count     All Obs        Obs

                        .           3        1.53     100.00
```

```
                  table 1: analysis of baseline factors


                       The UNIVARIATE Procedure
                           Variable:  wgtkg
                             RACEW = Yes

                               Moments

 N                          205    Sum Weights                205
 Mean                84.3104213    Sum Observations    17283.6364
 Std Deviation       17.7016357    Variance            313.347908
 Skewness            0.48340341    Kurtosis            0.22263762
 Uncorrected SS      1521113.64    Corrected SS        63922.9732
 Coeff Variation     20.9957861    Std Error Mean      1.23633587


                      Basic Statistical Measures

          Location                    Variability

      Mean     84.31042     Std Deviation            17.70164
      Median   82.27273     Variance                313.34791
      Mode     72.72727     Range                    97.72727
                            Interquartile Range      23.18182


                      Tests for Location: Mu0=0

         Test           -Statistic-    -----p Value------

         Student's t    t  68.19378    Pr > |t|    <.0001
         Sign           M     102.5    Pr >= |M|   <.0001
         Signed Rank    S   10557.5    Pr >= |S|   <.0001


                       Quantiles (Definition 5)

                       Quantile       Estimate

                       100% Max       148.6364
                       99%            126.8182
                       95%            115.9091
                       90%            106.8182
                       75% Q3          95.9091
                       50% Median      82.2727
                       25% Q1          72.7273
                       10%             60.4545
                       5%              56.3636
                       1%              52.2727
                       0% Min          50.9091
```

```
                    table 1: analysis of baseline factors


                           The UNIVARIATE Procedure
                               Variable:  wgtkg
                                 RACEW = Yes

                             Extreme Observations

                  ------Lowest-----        -----Highest-----

                     Value      Obs           Value      Obs

                   50.9091       70         125.000       12
                   52.2727      348         125.000      271
                   52.2727       23         126.818      388
                   53.1818      190         127.273      154
                   53.6364      313         148.636      239
```

```
                    table 1: analysis of baseline factors


                           The UNIVARIATE Procedure
                                Variable:  bmi
                                  RACEW = No

                                   Moments

            N                         191    Sum Weights                191
            Mean               30.6467432    Sum Observations    5853.52795
            Std Deviation      6.27370547    Variance            39.3593803
            Skewness           0.99454186    Kurtosis            1.36646596
            Uncorrected SS      186869.85    Corrected SS        7478.28226
            Coeff Variation    20.4710348    Std Error Mean       0.4539494


                          Basic Statistical Measures

                  Location                    Variability

              Mean     30.64674     Std Deviation            6.27371
              Median   29.49960     Variance                39.35938
              Mode     22.28403     Range                   36.69399
                                    Interquartile Range      7.85278

         NOTE: The mode displayed is the smallest of 4 modes with a count of 2.


                          Tests for Location: Mu0=0

                 Test           -Statistic-    -----p Value------

                 Student's t    t  67.51136    Pr > |t|    <.0001
                 Sign           M      95.5    Pr >= |M|   <.0001
                 Signed Rank    S      9168    Pr >= |S|   <.0001


                           Quantiles (Definition 5)

                            Quantile      Estimate

                            100% Max       56.0709
                            99%            49.7853
                            95%            41.2045
                            90%            38.8627
                            75% Q3         34.1851
                            50% Median     29.4996
                            25% Q1         26.3323
                            10%            23.6648
                            5%             22.2322
                            1%             19.9680
                            0% Min         19.3769
```

```
                    table 1: analysis of baseline factors

                         The UNIVARIATE Procedure
                              Variable:  bmi
                               RACEW = No

                           Extreme Observations

                   ------Lowest-----        -----Highest-----

                      Value      Obs           Value      Obs

                    19.3769       81         47.5516       18
                    19.9680      377         48.3796      202
                    21.0914      206         48.5581      105
                    21.3410      172         49.7853      341
                    21.6010      322         56.0709      368


                              Missing Values

                                         -----Percent Of-----
                 Missing                              Missing
                   Value       Count     All Obs          Obs

                       .           5        2.55       100.00
```

```
                     table 1: analysis of baseline factors


                            The UNIVARIATE Procedure
                                Variable:  bmi
                                 RACEW = Yes

                                   Moments

            N                          203    Sum Weights                203
            Mean                28.4650559    Sum Observations    5778.40635
            Std Deviation       5.64308585    Variance            31.8444179
            Skewness            1.13609849    Kurtosis            2.24236009
            Uncorrected SS      170915.232    Corrected SS        6432.57241
            Coeff Variation     19.8246083    Std Error Mean      0.39606698


                          Basic Statistical Measures

                 Location                    Variability

             Mean     28.46506     Std Deviation            5.64309
             Median   27.59877     Variance                31.84442
             Mode     29.08367     Range                   34.27125
                                   Interquartile Range      7.12627


                          Tests for Location: Mu0=0

                Test           -Statistic-    -----p Value------

                Student's t    t   71.8693    Pr > |t|    <.0001
                Sign           M     101.5    Pr >= |M|   <.0001
                Signed Rank    S     10353    Pr >= |S|   <.0001


                           Quantiles (Definition 5)

                            Quantile      Estimate

                            100% Max       53.2041
                            99%            45.8581
                            95%            37.8656
                            90%            35.4053
                            75% Q3         31.5395
                            50% Median     27.5988
                            25% Q1         24.4132
                            10%            21.8951
                            5%             21.0316
                            1%             19.4572
                            0% Min         18.9328
```

```
   table 1: analysis of baseline factors

          The UNIVARIATE Procedure
               Variable:  bmi
                RACEW = Yes

            Extreme Observations

 ------Lowest-----        -----Highest-----

    Value      Obs            Value      Obs

  18.9328      190          44.6683       42
  19.0708      103          45.7229      239
  19.4572      168          45.8581       12
  19.9037      237          47.8252      129
  20.1880       70          53.2041      329


               Missing Values

                       -----Percent Of-----
Missing                             Missing
  Value       Count     All Obs         Obs

     .            2        0.98      100.00
```

```
                    table 1: analysis of baseline factors

                          The UNIVARIATE Procedure
                      Variable:  ALT  (ALT result lab)
                                RACEW = No

                                  Moments

N                                194    Sum Weights                194
Mean                      71.0309278    Sum Observations         13780
Std Deviation             45.5106783    Variance            2071.22184
Skewness                  2.37506199    Kurtosis            9.59007239
Uncorrected SS               1378552    Corrected SS        399745.814
Coeff Variation            64.071637    Std Error Mean      3.26747628


                        Basic Statistical Measures

              Location                    Variability

          Mean     71.03093     Std Deviation            45.51068
          Median   59.00000     Variance                     2071
          Mode     40.00000     Range                   347.00000
                                Interquartile Range      50.00000


                        Tests for Location: Mu0=0

             Test           -Statistic-    -----p Value------

             Student's t    t  21.73877    Pr > |t|    <.0001
             Sign           M        97    Pr >= |M|   <.0001
             Signed Rank    S    9457.5    Pr >= |S|   <.0001


                         Quantiles (Definition 5)

                          Quantile      Estimate

                          100% Max           363
                          99%                238
                          95%                146
                          90%                126
                          75% Q3              90
                          50% Median          59
                          25% Q1              40
                          10%                 31
                          5%                  24
                          1%                  16
                          0% Min              16
```

```
                    table 1: analysis of baseline factors

                           The UNIVARIATE Procedure
                      Variable:  ALT  (ALT result lab)
                                 RACEW = No

                            Extreme Observations

                     ----Lowest----        ----Highest---

                     Value      Obs        Value      Obs

                        16      201          198      156
                        16       39          210      345
                        18      215          230       80
                        19      283          238      366
                        20       79          363      101


                               Missing Values

                                       -----Percent Of-----
                 Missing                              Missing
                   Value       Count     All Obs          Obs

                       .           1        0.51        50.00
                       B           1        0.51        50.00
                   Total           2        1.02       100.00
```

```
                    table 1: analysis of baseline factors


                          The UNIVARIATE Procedure
                      Variable:  ALT  (ALT result lab)
                                RACEW = Yes

                                  Moments

            N                          205    Sum Weights                205
            Mean                107.517073    Sum Observations         22041
            Std Deviation       90.6526226    Variance            8217.89799
            Skewness            2.08974237    Kurtosis            5.02759921
            Uncorrected SS         4046235    Corrected SS        1676451.19
            Coeff Variation     84.3146302    Std Error Mean       6.3314538


                         Basic Statistical Measures

                  Location                    Variability

               Mean     107.5171     Std Deviation            90.65262
               Median    74.0000     Variance                     8218
               Mode      54.0000     Range                   494.00000
                                     Interquartile Range      87.00000


                         Tests for Location: Mu0=0

                 Test           -Statistic-    -----p Value------

                 Student's t    t  16.98142    Pr > |t|    <.0001
                 Sign           M     102.5    Pr >= |M|   <.0001
                 Signed Rank    S   10557.5    Pr >= |S|   <.0001


                          Quantiles (Definition 5)

                           Quantile      Estimate

                           100% Max           508
                           99%                478
                           95%                281
                           90%                221
                           75% Q3             138
                           50% Median          74
                           25% Q1              51
                           10%                 36
                           5%                  28
                           1%                  21
                           0% Min              14
```

```
table 1: analysis of baseline factors


        The UNIVARIATE Procedure
    Variable:  ALT  (ALT result lab)
              RACEW = Yes

         Extreme Observations

 ----Lowest----        ----Highest---

 Value      Obs        Value      Obs

    14      263          409      238
    20      129          428      299
    21       70          478      302
    23      281          505       35
    23      164          508      378
```

```
                        table 1: analysis of baseline factors


                              The UNIVARIATE Procedure
                          Variable:  AST  (AST result lab)
                                    RACEW = No

                                      Moments

            N                          194    Sum Weights                194
            Mean                59.0360825    Sum Observations         11453
            Std Deviation        41.079775    Variance            1687.54791
            Skewness            3.46092158    Kurtosis            18.1765964
            Uncorrected SS         1001837    Corrected SS        325696.747
            Coeff Variation     69.5841819    Std Error Mean      2.94935596


                             Basic Statistical Measures

                   Location                    Variability

               Mean     59.03608     Std Deviation            41.07978
               Median   51.00000     Variance                     1688
               Mode     24.00000     Range                   345.00000
                                     Interquartile Range      36.00000

         NOTE: The mode displayed is the smallest of 3 modes with a count of 6.


                             Tests for Location: Mu0=0

                  Test           -Statistic-    -----p Value------

                  Student's t    t   20.0166    Pr > |t|    <.0001
                  Sign           M        97    Pr >= |M|   <.0001
                  Signed Rank    S    9457.5    Pr >= |S|   <.0001


                              Quantiles (Definition 5)

                                Quantile      Estimate

                                100% Max           363
                                99%                234
                                95%                126
                                90%                 95
                                75% Q3              69
                                50% Median          51
                                25% Q1              33
                                10%                 25
                                5%                  24
                                1%                  19
                                0% Min              18
```

```
                 table 1: analysis of baseline factors


                      The UNIVARIATE Procedure
                   Variable:  AST  (AST result lab)
                             RACEW = No

                        Extreme Observations

                  ----Lowest----        ----Highest---

                  Value      Obs        Value      Obs

                     18       39          195      197
                     19      283          203       80
                     20      195          208      360
                     21      277          234      345
                     23      398          363      101


                           Missing Values

                                    -----Percent Of-----
              Missing                             Missing
                Value       Count     All Obs         Obs

                    .           1        0.51       50.00
                    B           1        0.51       50.00
                Total           2        1.02      100.00
```

```
                    table 1: analysis of baseline factors


                          The UNIVARIATE Procedure
                      Variable:  AST  (AST result lab)
                                RACEW = Yes

                                  Moments

            N                          205    Sum Weights                205
            Mean                72.9609756    Sum Observations         14957
            Std Deviation        59.382595    Variance            3526.29259
            Skewness            2.66525749    Kurtosis            9.97105001
            Uncorrected SS         1810641    Corrected SS        719363.688
            Coeff Variation     81.3895298    Std Error Mean      4.14746034


                         Basic Statistical Measures

                 Location                    Variability

             Mean     72.96098     Std Deviation            59.38259
             Median   52.00000     Variance                     3526
             Mode     36.00000     Range                   439.00000
                                   Interquartile Range      50.00000


                         Tests for Location: Mu0=0

                Test           -Statistic-    -----p Value------

                Student's t    t  17.59172    Pr > |t|    <.0001
                Sign           M     102.5    Pr >= |M|   <.0001
                Signed Rank    S   10557.5    Pr >= |S|   <.0001


                          Quantiles (Definition 5)

                           Quantile      Estimate

                           100% Max           455
                           99%                275
                           95%                207
                           90%                142
                           75% Q3              87
                           50% Median          52
                           25% Q1              37
                           10%                 27
                           5%                  24
                           1%                  18
                           0% Min              16
```

```
                    table 1: analysis of baseline factors

                        The UNIVARIATE Procedure
                    Variable:  AST  (AST result lab)
                              RACEW = Yes

                          Extreme Observations

                     ----Lowest----        ----Highest---

                     Value      Obs        Value      Obs

                        16       70          263      378
                        17      251          269      238
                        18      263          275       68
                        18      129          283      192
                        20      177          455      326
```

```
                  table 1: analysis of baseline factors


                       The UNIVARIATE Procedure
                  Variable:  ALB  (Albumin result lab)
                              RACEW = No

                                Moments

N                          139    Sum Weights                139
Mean                3.99208633    Sum Observations         554.9
Std Deviation       0.38861791    Variance            0.15102388
Skewness            -0.5865309    Kurtosis            2.12718791
Uncorrected SS         2236.05    Corrected SS         20.841295
Coeff Variation     9.73470695    Std Error Mean      0.03296214


                      Basic Statistical Measures

            Location                    Variability

        Mean     3.992086     Std Deviation            0.38862
        Median   4.000000     Variance                 0.15102
        Mode     4.000000     Range                    2.70000
                              Interquartile Range      0.50000


                      Tests for Location: Mu0=0

            Test           -Statistic-    -----p Value------

            Student's t    t  121.1113    Pr > |t|    <.0001
            Sign           M      69.5    Pr >= |M|   <.0001
            Signed Rank    S      4865    Pr >= |S|   <.0001


                       Quantiles (Definition 5)

                        Quantile      Estimate

                        100% Max           5.0
                        99%                4.8
                        95%                4.6
                        90%                4.4
                        75% Q3             4.2
                        50% Median         4.0
                        25% Q1             3.7
                        10%                3.5
                        5%                 3.4
                        1%                 2.9
                        0% Min             2.3
```

```
                   table 1: analysis of baseline factors

                          The UNIVARIATE Procedure
                    Variable:  ALB  (Albumin result lab)
                                RACEW = No

                           Extreme Observations

                    ----Lowest----        ----Highest---

                    Value      Obs        Value      Obs

                      2.3       83          4.7       77
                      2.9      340          4.7       81
                      3.1      396          4.7      180
                      3.1      314          4.8      291
                      3.4      394          5.0      182


                              Missing Values

                                        -----Percent Of-----
                Missing                               Missing
                  Value       Count     All Obs           Obs

                      .           1        0.51          1.75
                      B          56       28.57         98.25
                  Total          57       29.08        100.00
```

```
                    table 1: analysis of baseline factors


                          The UNIVARIATE Procedure
                     Variable:  ALB  (Albumin result lab)
                                RACEW = Yes

                                  Moments

        N                          143    Sum Weights                143
        Mean                4.17342657    Sum Observations         596.8
        Std Deviation       0.32565234    Variance            0.10604944
        Skewness              -0.18588    Kurtosis            -0.1270047
        Uncorrected SS         2505.76    Corrected SS         15.059021
        Coeff Variation     7.80299664    Std Error Mean      0.02723242


                         Basic Statistical Measures

               Location                    Variability

           Mean     4.173427     Std Deviation            0.32565
           Median   4.200000     Variance                 0.10605
           Mode     4.100000     Range                    1.70000
                                 Interquartile Range      0.40000

    NOTE: The mode displayed is the smallest of 2 modes with a count of 19.


                         Tests for Location: Mu0=0

              Test           -Statistic-    -----p Value------

              Student's t    t  153.2522    Pr > |t|    <.0001
              Sign           M      71.5    Pr >= |M|   <.0001
              Signed Rank    S      5148    Pr >= |S|   <.0001


                          Quantiles (Definition 5)

                           Quantile      Estimate

                           100% Max           5.0
                           99%                4.9
                           95%                4.6
                           90%                4.6
                           75% Q3             4.4
                           50% Median         4.2
                           25% Q1             4.0
                           10%                3.8
                           5%                 3.7
                           1%                 3.3
                           0% Min             3.3
```

```
                table 1: analysis of baseline factors


                       The UNIVARIATE Procedure
                Va riable:  ALB  (Albumin result lab)
                             RACEW = Yes

                        Extreme Observations

                 ----Lowest----        ----Highest---

                Va lue      Obs        Value      Obs

                   3.3      303          4.7      388
                   3.3      142          4.8      114
                   3.5      168          4.8      347
                   3.5      159          4.9      335
                   3.5      129          5.0      331


                           Missing Values

                                    -----Percent Of-----
             Missing                             Missing
               Value       Count     All Obs         Obs

                   B          62       30.24      100.00
```

```
             table 1: analysis of baseline factors


                    The UNIVARIATE Procedure
         Variable:  TBILI  (Total bilirubin lab result)
                           RACEW = No

                            Moments

N                              194    Sum Weights                194
Mean                     0.6314433    Sum Observations         122.5
Std Deviation           0.31042505    Variance            0.09636371
Skewness                1.72084376    Kurtosis            5.67232557
Uncorrected SS               95.95    Corrected SS        18.5981959
Coeff Variation          49.161191    Std Error Mean      0.02228722


                   Basic Statistical Measures

         Location                    Variability

     Mean     0.631443     Std Deviation            0.31043
     Median   0.600000     Variance                 0.09636
     Mode     0.400000     Range                    2.30000
                           Interquartile Range      0.40000


                   Tests for Location: Mu0=0

        Test           -Statistic-    -----p Value------

        Student's t    t  28.33208    Pr > |t|    <.0001
        Sign           M        97    Pr >= |M|   <.0001
        Signed Rank    S    9457.5    Pr >= |S|   <.0001


                    Quantiles (Definition 5)

                     Quantile      Estimate

                     100% Max           2.4
                     99%                1.7
                     95%                1.2
                     90%                0.9
                     75% Q3             0.8
                     50% Median         0.6
                     25% Q1             0.4
                     10%                0.3
                     5%                 0.3
                     1%                 0.2
                     0% Min             0.1
```

```
                 table 1: analysis of baseline factors


                       The UNIVARIATE Procedure
          Variable:  TBILI  (Total bilirubin lab result)
                              RACEW = No

                         Extreme Observations

                  ----Lowest----        ----Highest---

                  Value      Obs        Value      Obs

                    0.1      323          1.4      280
                    0.2      399          1.5       51
                    0.2      398          1.5      125
                    0.2      362          1.7      345
                    0.2      205          2.4      314


                            Missing Values

                                     -----Percent Of-----
           Missing                                Missing
             Value       Count     All Obs            Obs

                 .           1        0.51          50.00
                 B           1        0.51          50.00
             Total           2        1.02         100.00
```

```
                      table 1: analysis of baseline factors


                           The UNIVARIATE Procedure
                  Variable:  TBILI  (Total bilirubin lab result)
                                 RACEW = Yes

                                   Moments

         N                         204    Sum Weights                204
         Mean               0.74264706    Sum Observations         151.5
         Std Deviation      0.38146058    Variance            0.14551217
         Skewness           2.27308677    Kurtosis            9.31999064
         Uncorrected SS         142.05    Corrected SS        29.5389706
         Coeff Variation    51.3649883    Std Error Mean      0.02670758


                          Basic Statistical Measures

                Location                    Variability

            Mean     0.742647     Std Deviation            0.38146
            Median   0.700000     Variance                 0.14551
            Mode     0.700000     Range                    2.80000
                                  Interquartile Range      0.40000


                          Tests for Location: Mu0=0

               Test           -Statistic-    -----p Value------

               Student's t    t   27.8066    Pr > |t|    <.0001
               Sign           M       102    Pr >= |M|   <.0001
               Signed Rank    S     10455    Pr >= |S|   <.0001


                           Quantiles (Definition 5)

                            Quantile      Estimate

                            100% Max           2.9
                            99%                1.9
                            95%                1.4
                            90%                1.2
                            75% Q3             0.9
                            50% Median         0.7
                            25% Q1             0.5
                            10%                0.4
                            5%                 0.3
                            1%                 0.2
                            0% Min             0.1
```

```
                The UNIVARIATE Procedure
       Variable:  TBILI  (Total bilirubin lab result)
                     RACEW = Yes

                 Extreme Observations

          ----Lowest----        ----Highest---

          Value      Obs        Value      Obs

            0.1      210          1.7       66
            0.2      348          1.9       32
            0.2      320          1.9      378
            0.2      306          2.9       68
            0.2      303          2.9      395


                    Missing Values

                             -----Percent Of-----
        Missing                              Missing
          Value       Count      All Obs          Obs

              B           1         0.49       100.00
```

```
                     table 1: analysis of baseline factors


                          The UNIVARIATE Procedure
                      Variable:  HGB  (Hemoglobin lab)
                                 RACEW = No

                                  Moments

         N                         193    Sum Weights                193
         Mean               14.3165803    Sum Observations        2763.1
         Std Deviation      1.31934118    Variance            1.74066116
         Skewness           0.00938519    Kurtosis            0.32578536
         Uncorrected SS       39892.35    Corrected SS        334.206943
         Coeff Variation    9.21547713    Std Error Mean      0.09496826


                         Basic Statistical Measures

             Location                    Variability

          Mean     14.31658     Std Deviation            1.31934
          Median   14.30000     Variance                 1.74066
          Mode     13.70000     Range                    7.30000
                                Interquartile Range      1.60000

     NOTE: The mode displayed is the smallest of 2 modes with a count of 11.


                         Tests for Location: Mu0=0

             Test           -Statistic-    -----p Value------

             Student's t    t  150.7512    Pr > |t|    <.0001
             Sign           M      96.5    Pr >= |M|   <.0001
             Signed Rank    S    9360.5    Pr >= |S|   <.0001


                          Quantiles (Definition 5)

                           Quantile      Estimate

                           100% Max          18.0
                           99%               17.8
                           95%               16.5
                           90%               16.0
                           75% Q3            15.1
                           50% Median        14.3
                           25% Q1            13.5
                           10%               12.7
                           5%                11.9
                           1%                10.8
                           0% Min            10.7
```

```
          The UNIVARIATE Procedure
      Variable:  HGB  (Hemoglobin lab)
                 RACEW = No

            Extreme Observations

    ----Lowest----        ----Highest---

    Value      Obs        Value      Obs

     10.7       83         17.1      393
     10.8      340         17.2        2
     11.3      195         17.6      246
     11.4      205         17.8      145
     11.6      197         18.0      182


               Missing Values

                       -----Percent Of-----
Missing                             Missing
  Value       Count     All Obs         Obs

      .           1        0.51       33.33
     B            2        1.02       66.67
  Total           3        1.53      100.00
```

```
                      table 1: analysis of baseline factors


                             The UNIVARIATE Procedure
                         Variable:  HGB  (Hemoglobin lab)
                                   RACEW = Yes

                                     Moments

N                             202    Sum Weights                202
 Mean                  14.9386139    Sum Observations        3017.6
 Std Deviation         1.29967239    Variance            1.68914832
 Skewness              -0.0400601    Kurtosis            -0.4356941
 Uncorrected SS          45418.28    Corrected SS        339.518812
 Coeff Variation       8.70008691    Std Error Mean      0.09144463


                            Basic Statistical Measures

                  Location                    Variability

              Mean     14.93861     Std Deviation             1.29967
              Median   15.05000     Variance                  1.68915
              Mode     15.50000     Range                     5.90000
                                    Interquartile Range       2.00000


                            Tests for Location: Mu0=0

                 Test           -Statistic-    -----p Value------

                 Student's t    t  163.3624    Pr > |t|    <.0001
                 Sign           M       101    Pr >= |M|   <.0001
                 Signed Rank    S   10251.5    Pr >= |S|   <.0001


                             Quantiles (Definition 5)

                              Quantile      Estimate

                              100% Max         17.70
                              99%              17.60
                              95%              17.20
                              90%              16.60
                              75% Q3           15.90
                              50% Median       15.05
                              25% Q1           13.90
                              10%              13.30
                              5%               12.80
                              1%               11.90
                              0% Min           11.80
```

```
                    table 1: analysis of baseline factors


                           The UNIVARIATE Procedure
                        Variable:  HGB  (Hemoglobin lab)
                                 RACEW = Yes

                             Extreme Observations

                     ----Lowest----        ----Highest---

                     Value      Obs        Value      Obs

                      11.8      303         17.5      311
                      11.9      198         17.5      387
                      11.9      168         17.6      335
                      12.1      113         17.7      124
                      12.4      329         17.7      236


                                Missing Values

                                        -----Percent Of-----
                 Missing                             Missing
                   Value       Count     All Obs         Obs

                       B           2        0.98       66.67
                       F           1        0.49       33.33
                   Total           3        1.46      100.00
```

```
                     table 1: analysis of baseline factors


                           The UNIVARIATE Procedure
                    Variable:  WBC  (White blood cells lab)
                                  RACEW = No

                                    Moments

            N                          193    Sum Weights                193
            Mean                6.13782383    Sum Observations        1184.6
            Std Deviation       2.09790804    Variance            4.40121816
            Skewness            0.79840598    Kurtosis             0.1966282
            Uncorrected SS          8115.9    Corrected SS        845.033886
            Coeff Variation     34.1799977    Std Error Mean      0.15101073


                          Basic Statistical Measures

                  Location                    Variability

              Mean     6.137824     Std Deviation            2.09791
              Median   5.800000     Variance                 4.40122
              Mode     4.800000     Range                   10.00000
                                    Interquartile Range      2.70000


                          Tests for Location: Mu0=0

                Test           -Statistic-    -----p Value------

                Student's t    t  40.64495    Pr > |t|    <.0001
                Sign           M      96.5    Pr >= |M|   <.0001
                Signed Rank    S    9360.5    Pr >= |S|   <.0001


                           Quantiles (Definition 5)

                            Quantile      Estimate

                            100% Max          12.7
                            99%               12.1
                            95%               10.5
                            90%                9.4
                            75% Q3             7.4
                            50% Median         5.8
                            25% Q1             4.7
                            10%                3.7
                            5%                 3.3
                            1%                 2.7
                            0% Min             2.7
```

```
   table 1: analysis of baseline factors


          The UNIVARIATE Procedure
  Variable:  WBC  (White blood cells lab)
                 RACEW = No

            Extreme Observations

    ----Lowest----        ----Highest---

    Value      Obs        Value      Obs

      2.7      314         11.0      141
      2.7      293         11.3       80
      2.7      206         11.3      183
      2.8      349         12.1      362
      2.9      280         12.7       96


               Missing Values

                       -----Percent Of-----
Missing                            Missing
  Value       Count     All Obs        Obs

      .           1        0.51      33.33
      B           2        1.02      66.67
  Total           3        1.53     100.00
```

```
                    table 1: analysis of baseline factors


                         The UNIVARIATE Procedure
                   Variable:  WBC  (White blood cells lab)
                              RACEW = Yes

                                 Moments

    N                         202    Sum Weights                202
    Mean               6.41237624    Sum Observations        1295.3
    Std Deviation      2.00784414    Variance            4.03143811
    Skewness           1.22975499    Kurtosis            2.42528399
    Uncorrected SS        9116.27    Corrected SS        810.319059
    Coeff Variation     31.312014    Std Error Mean      0.14127142


                        Basic Statistical Measures

              Location                    Variability

         Mean     6.412376     Std Deviation            2.00784
         Median   6.200000     Variance                 4.03144
         Mode     4.900000     Range                   12.50000
                               Interquartile Range      2.50000


                        Tests for Location: Mu0=0

            Test           -Statistic-    -----p Value------

            Student's t    t  45.39047    Pr > |t|    <.0001
            Sign           M       101    Pr >= |M|   <.0001
            Signed Rank    S   10251.5    Pr >= |S|   <.0001


                         Quantiles (Definition 5)

                          Quantile      Estimate

                          100% Max          14.5
                          99%               13.6
                          95%               10.2
                          90%                8.7
                          75% Q3             7.4
                          50% Median         6.2
                          25% Q1             4.9
                          10%                4.4
                          5%                 3.9
                          1%                 3.3
                          0% Min             2.0
```

```
                 table 1: analysis of baseline factors


                        The UNIVARIATE Procedure
                 Variable:  WBC  (White blood cells lab)
                              RACEW = Yes

                          Extreme Observations

                   ----Lowest----        ----Highest---

                   Value      Obs        Value      Obs

                     2.0      149         11.9      200
                     3.0       88         12.5      194
                     3.3      119         13.6      306
                     3.5      208         13.9      387
                     3.6      237         14.5      342


                             Missing Values

                                        -----Percent Of-----
              Missing                                Missing
                Value       Count     All Obs           Obs

                    B           2        0.98         66.67
                    F           1        0.49         33.33
                Total           3        1.46        100.00
```

```
                    table 1: analysis of baseline factors


                            The UNIVARIATE Procedure
                             Variable:  log10vload
                                  RACEW = No

                                    Moments

N                         194    Sum Weights                194
Mean               6.23473169    Sum Observations    1209.53795
Std Deviation      0.66978945    Variance            0.44861791
Skewness           -0.7136958    Kurtosis            0.84701483
Uncorrected SS     7627.72783    Corrected SS        86.5832567
Coeff Variation    10.7428753    Std Error Mean      0.04808808


                           Basic Statistical Measures

             Location                    Variability

         Mean     6.234732     Std Deviation            0.66979
         Median   6.393574     Variance                 0.44862
         Mode     6.563481     Range                    4.23946
                               Interquartile Range      1.07918


                           Tests for Location: Mu0=0

            Test           -Statistic-    -----p Value------

            Student's t    t  129.6523    Pr > |t|    <.0001
            Sign           M        97    Pr >= |M|   <.0001
            Signed Rank    S    9457.5    Pr >= |S|   <.0001


                            Quantiles (Definition 5)

                             Quantile      Estimate

                             100% Max       7.48001
                             99%            7.38917
                             95%            7.15229
                             90%            7.01284
                             75% Q3         6.72754
                             50% Median     6.39357
                             25% Q1         5.64836
                             10%            5.32428
                             5%             5.11394
                             1%             4.76790
                             0% Min         3.24055
```

```
   table 1: analysis of baseline factors


          The UNIVARIATE Procedure
           Variable:  log10vload
                 RACEW = No

            Extreme Observations

 ------Lowest-----         -----Highest-----

    Value      Obs             Value      Obs

  3.24055      229           7.26717      332
  4.76790      172           7.29003      257
  4.85552       86           7.34830      272
  4.88480       52           7.38917      397
  4.98722      375           7.48001      144


               Missing Values

                       -----Percent Of-----
Missing                             Missing
  Value       Count     All Obs         Obs

      .           2        1.02      100.00
```

```
                  table 1: analysis of baseline factors


                         The UNIVARIATE Procedure
                          Variable:  log10vload
                               RACEW = Yes

                                 Moments

N                         205    Sum Weights                205
Mean               6.29757466    Sum Observations    1291.00281
Std Deviation      0.75959726    Variance              0.576988
Skewness           -1.1676952    Kurtosis            1.63028543
Uncorrected SS     8247.89211    Corrected SS        117.705552
Coeff Variation    12.0617428    Std Error Mean      0.05305257


                        Basic Statistical Measures

          Location                    Variability

     Mean     6.297575     Std Deviation            0.75960
     Median   6.542825     Variance                 0.57699
     Mode     5.432969     Range                    4.53630
                           Interquartile Range      1.14972

NOTE: The mode displayed is the smallest of 12 modes with a count of 2.


                        Tests for Location: Mu0=0

        Test           -Statistic-    -----p Value------

        Student's t    t  118.7044    Pr > |t|    <.0001
        Sign           M     102.5    Pr >= |M|   <.0001
        Signed Rank    S   10557.5    Pr >= |S|   <.0001


                         Quantiles (Definition 5)

                          Quantile      Estimate

                          100% Max       7.67302
                          99%            7.46240
                          95%            7.19312
                          90%            7.07188
                          75% Q3         6.80003
                          50% Median     6.54283
                          25% Q1         5.65031
                          10%            5.32222
                          5%             4.99167
                          1%             3.95617
                          0% Min         3.13672
```

```
                table 1: analysis of baseline factors


                       The UNIVARIATE Procedure
                        Variable:  log10vload
                             RACEW = Yes

                         Extreme Observations

                ------Lowest-----        -----Highest-----

                   Value      Obs           Value      Obs

                 3.13672      163         7.32015      369
                 3.58771      308         7.35984      278
                 3.95617      186         7.46240      167
                 4.10380      129         7.56703      347
                 4.35793      164         7.67302       21
```

```
                  table 1: analysis of baseline factors

                          The NPAR1WAY Procedure

              Wilcoxon Scores (Rank Sums) for Variable age
                      Classified by Variable RACEW

                          Sum of      Expected       Std Dev          Mean
RACEW        N          Scores      Under H0      Under H0         Score
---------------------------------------------------------------------
Yes        205         39091.0       41205.0    1160.15899    190.687805
No         196         41510.0       39396.0    1160.15899    211.785714

                   Average scores were used for ties.


                        Wilcoxon Two-Sample Test

                    Statistic             41510.0000

                    Normal Approximation
                    Z                         1.8217
                    One-Sided Pr >  Z         0.0342
                    Two-Sided Pr > |Z|        0.0685

                    t Approximation
                    One-Sided Pr >  Z         0.0346
                    Two-Sided Pr > |Z|        0.0692

              Z includes a continuity correction of 0.5.


                          Kruskal-Wallis Test

                    Chi-Square                3.3203
                    DF                             1
                    Pr > Chi-Square           0.0684
```

```
                          The NPAR1WAY Procedure

             Wilcoxon Scores (Rank Sums) for Variable wgtkg
                       Classified by Variable RACEW

                          Sum of      Expected       Std Dev          Mean
RACEW        N          Scores      Under H0      Under H0         Score
--------------------------------------------------------------------
Yes        205         37201.0      40897.50    1146.89657    181.468293
No         193         42200.0      38503.50    1146.89657    218.652850

                   Average scores were used for ties.


                        Wilcoxon Two-Sample Test

                   Statistic             42200.0000

                   Normal Approximation
                   Z                         3.2226
                   One-Sided Pr >  Z         0.0006
                   Two-Sided Pr > |Z|        0.0013

                   t Approximation
                   One-Sided Pr >  Z         0.0007
                   Two-Sided Pr > |Z|        0.0014

              Z includes a continuity correction of 0.5.


                          Kruskal-Wallis Test

                   Chi-Square               10.3880
                   DF                             1
                   Pr > Chi-Square           0.0013
```

```
                    table 1: analysis of baseline factors


                           The NPAR1WAY Procedure

                Wilcoxon Scores (Rank Sums) for Variable bmi
                        Classified by Variable RACEW

                      Sum of      Expected       Std Dev          Mean
RACEW        N        Scores      Under H0      Under H0         Score
----------------------------------------------------------------------
Yes        203       35995.0      40092.50    1129.72327    177.315271
No         191       41820.0      37722.50    1129.72327    218.952880

                   Average scores were used for ties.


                         Wilcoxon Two-Sample Test

                       Statistic            41820.0000

                       Normal Approximation
                      Z                         3.6266
                      One-Sided Pr >  Z         0.0001
                      Two-Sided Pr > |Z|        0.0003

                       t Approximation
                      One-Sided Pr >  Z         0.0002
                      Two-Sided Pr > |Z|        0.0003

              Z includes a continuity correction of 0.5.


                           Kruskal-Wallis Test

                      Chi-Square               13.1551
                      DF                             1
                      Pr > Chi-Square           0.0003
```

```
                    ta ble 1: analysis of baseline factors


                             The NPAR1WAY Procedure

                  Wilcoxon Scores (Rank Sums) for Variable ALT
                          Classified by Variable RACEW

                         Sum of      Expected       Std Dev          Mean
RACEW        N           Scores      Under H0      Under H0         Score
--------------------------------------------------------------------------
Yes        205         45846.50       41000.0    1151.28180    223.641463
No         194         33953.50       38800.0    1151.28180    175.018041

                       Average scores were used for ties.


                           Wilcoxon Two -Sample Test

                         Statistic             33953.5000

                         Normal  Approximation
                        Z                         -4.2092
                         One -Sided Pr <  Z        <.0001
                         Two -Sided Pr > |Z|       <.0001

                         t Approximation
                         One -Sided Pr <  Z        <.0001
                         Two -Sided Pr > |Z|       <.0001

                  Z includes a continuity correction of 0.5.


                             Kruskal -Wallis Test

                         Chi -Square              17.7212
                         DF                             1
                         Pr > Chi -Square          <.0001
```

```
                   table 1: analysis of baseline factors


                           The NPAR1WAY Procedure

                Wilcoxon Scores (Rank Sums) for Variable AST
                        Classified by Variable RACEW

                           Sum of      Expected       Std Dev          Mean
RACEW         N            Scores      Under H0      Under H0         Score
---------------------------------------------------------------------------
Yes         205          43197.50       41000.0    1151.25407    210.719512
No          194          36602.50       38800.0    1151.25407    188.672680

                     Average scores were used for ties.


                         Wilcoxon Two-Sample Test

                       Statistic             36602.5000

                       Normal Approximation
                      Z                         -1.9084
                      One-Sided Pr <  Z          0.0282
                      Two-Sided Pr > |Z|         0.0563

                       t Approximation
                      One-Sided Pr <  Z          0.0285
                      Two-Sided Pr > |Z|         0.0571

                 Z includes a continuity correction of 0.5.


                           Kruskal-Wallis Test

                      Chi-Square                 3.6435
                      DF                              1
                      Pr > Chi-Square            0.0563
```

```
                    table 1: analysis of baseline factors


                            The NPAR1WAY Procedure

                Wilcoxon Scores (Rank Sums) for Variable ALB
                        Classified by Variable RACEW

                          Sum of      Expected       Std Dev          Mean
RACEW        N            Scores      Under H0      Under H0         Score
---------------------------------------------------------------------------
Yes        143          23040.50      20234.50    681.850460    161.122378
No         139          16862.50      19668.50    681.850460    121.312950

                     Average scores were used for ties.


                         Wilcoxon Two -Sample Test

                        Statistic             16862.5000

                        Normal Approximation
                       Z                         -4.1145
                       One -Sided Pr <  Z         <.0001
                       Two -Sided Pr > |Z|        <.0001

                        t Approximation
                       One -Sided Pr <  Z         <.0001
                       Two -Sided Pr > |Z|        <.0001

                Z includes a continuity correction of 0.5.


                           Kruskal-Wallis Test

                       Chi -Square               16.9355
                       DF                              1
                       Pr > Chi -Square           <.0001
```

```
                  table 1: analysis of baseline factors


                         The NPAR1WAY Procedure

              Wilcoxon Scores (Rank Sums) for Variable TBILI
                      Classified by Variable RACEW

                       Sum of      Expected       Std Dev          Mean
RACEW        N         Scores      Under H0      Under H0         Score
--------------------------------------------------------------------
Yes        204       44525.50       40698.0    1139.24138    218.262255
No         194       34875.50       38703.0    1139.24138    179.770619

                   Average scores were used for ties.


                       Wilcoxon Two-Sample Test

                   Statistic            34875.5000

                   Normal Approximation
                  Z                        -3.3593
                   One-Sided Pr <  Z        0.0004
                   Two-Sided Pr > |Z|       0.0008

                   t Approximation
                   One-Sided Pr <  Z        0.0004
                   Two-Sided Pr > |Z|       0.0009

              Z includes a continuity correction of 0.5.


                         Kruskal-Wallis Test

                   Chi-Square              11.2875
                   DF                            1
                   Pr > Chi-Square          0.0008
```

```
                    table 1: analysis of baseline factors


                           The NPAR1WAY Procedure

                Wilcoxon Scores (Rank Sums) for Variable HGB
                        Classified by Variable RACEW

                          Sum of      Expected       Std Dev          Mean
RACEW         N           Scores      Under H0      Under H0         Score
--------------------------------------------------------------------------
Yes         202          45168.0       39996.0    1133.85268    223.603960
No          193          33042.0       38214.0    1133.85268    171.202073

                     Average scores were used for ties.


                          Wilcoxon Two-Sample Test

                        Statistic             33042.0000

                        Normal Approximation
                       Z                         -4.5610
                        One-Sided Pr <  Z         <.0001
                        Two-Sided Pr > |Z|        <.0001

                        t Approximation
                        One-Sided Pr <  Z         <.0001
                        Two-Sided Pr > |Z|        <.0001

               Z includes a continuity correction of 0.5.


                            Kruskal-Wallis Test

                        Chi-Square               20.8067
                        DF                             1
                        Pr > Chi-Square           <.0001
```

```
                    table 1: analysis of baseline factors


                            The NPAR1WAY Procedure

               Wilcoxon Scores (Rank Sums) for Variable WBC
                       Classified by Variable RACEW

                          Sum of      Expected       Std Dev          Mean
RACEW        N           Scores      Under H0      Under H0         Score
--------------------------------------------------------------------------
Yes        202         41913.50       39996.0    1134.03057    207.492574
No         193         36296.50       38214.0    1134.03057    188.064767

                     Average scores were used for ties.


                         Wilcoxon Two-Sample Test

                       Statistic             36296.5000

                       Normal Approximation
                      Z                         -1.6904
                      One-Sided Pr <  Z          0.0455
                      Two-Sided Pr > |Z|         0.0909

                       t Approximation
                      One-Sided Pr <  Z          0.0459
                      Two-Sided Pr > |Z|         0.0917

                  Z includes a continuity correction of 0.5.


                           Kruskal-Wallis Test

                      Chi-Square                 2.8590
                      DF                              1
                      Pr > Chi-Square            0.0909
```

```
                    table 1: analysis of baseline factors


                            The NPAR1WAY Procedure

             Wilcoxon Scores (Rank Sums) for Variable log10vload
                        Classified by Variable RACEW

                         Sum of      Expected       Std Dev          Mean
RACEW        N           Scores      Under H0      Under H0         Score
---------------------------------------------------------------------
Yes        205          42924.0       41000.0    1151.37262    209.385366
No         194          36876.0       38800.0    1151.37262    190.082474

                     Average scores were used for ties.


                         Wilcoxon Two-Sample Test

                       Statistic            36876.0000

                       Normal Approximation
                      Z                        -1.6706
                      One-Sided Pr <  Z         0.0474
                      Two-Sided Pr > |Z|        0.0948

                       t Approximation
                      One-Sided Pr <  Z         0.0478
                      Two-Sided Pr > |Z|        0.0956

                 Z includes a continuity correction of 0.5.


                           Kruskal-Wallis Test

                      Chi-Square                2.7924
                      DF                             1
                      Pr > Chi-Square           0.0947
```

Attachment 5

SAS 9.1 Output for programming code submitted for the replication of selected results in Table 3 “Relationship Between SVR and Each Variable of Interest in Univariable and Multivariable Analysis” Conjeevaram et.al. 2006

```
         replication of selected results in table 3: SVR<=>variables of interest


                                    The GENMOD Procedure

                                     Model Information

                          Data Set                 WORK.ALL24_FU2
                          Distribution                   Binomial
                          Link Function                       Log
                          Dependent Variable                  svr


                        Number of Observations Read         401
                        Number of Observations Used         401
                        Number of Events                    160
                        Number of Trials                    401


                                      Response Profile

                               Ordered                  Total
                                 Value     svr      Frequency

                                     1     1              160
                                     2     0              241

PROC GENMOD is modeling the probability that svr='1'.


                                   Par ameter Information

                               Parameter       Effect

                               Prm1            Intercept
                               Prm2            RACEW


                          Criteria For Assessing Goodness Of Fit

                 Criterion                 DF           Value        Value/DF

                 Deviance                 399        512.5915          1.2847
                 Scaled Deviance          399        512.5915          1.2847
                 Pearson Chi-Square       399        401.0000          1.0050
                 Scaled Pearson X2        399        401.0000          1.0050
                 Log Likelihood                     -256.2958


          Algorithm converged.
```

```
              replication of selected results in table 3: SVR<=>variables of interest

                                         The GENMOD Procedure

                                    Analysis Of Parameter Estimates

                                           Standard     Wald 95% Confidence       Chi-
     Parameter    DF    Estimate        Error           Limits             Square    Pr > ChiSq

     Intercept     1     -1.3078       0.1173      -1.5378     -1.0779      124.25        <.0001
     RACEW         1      0.6576       0.1350       0.3930      0.9223       23.72        <.0001
     Scale         0      1.0000       0.0000       1.0000      1.0000

NOTE: The scale parameter was held fixed.


                                   LR Statistics For Type 3 Analysis

                                                     Chi -
                              Source           DF     Square    Pr > ChiSq

                              RACEW             1      26.84        <.0001


                                       Contrast Estimate Results

                                      Standard                                        Chi-
Label                    Estimate        Error      Alpha      Confidence Limits      Square    Pr > ChiSq

Beta racew                 0.6576       0.1350       0.05      0.3930       0.9223     23.72        <.0001
Exp(Beta racew)            1.9302       0.2606       0.05      1.4814       2.5151
```

```
          replication of selected results in table 3: SVR<=>variables of interest

                                    The GENMOD Procedure

                                     Model Information

                          Data Set                WORK.ALL24_FU2
                          Distribution                  Binomial
                          Link Function                      Log
                          Dependent Variable                 svr


                        Number of Observations Read         401
                        Number of Observations Used         401
                        Number of Events                    160
                        Number of Trials                    401


                                      Response Profile

                               Ordered               Total
                                 Value    svr    Frequency

                                     1    1            160
                                     2    0            241

ROC GENMOD is modeling the probability that svr='1'.


                                   Parameter Information

                               Parameter       Effect

                               Prm1            Intercept
                               Prm2            male


                           Criteria For Assessing Goodness Of Fit

                 Criterion                 DF           Value        Value/DF

                 Deviance                 399        533.7882          1.3378
                 Scaled Deviance          399        533.7882          1.3378
                 Pearson Chi-Square       399        401.0000          1.0050
                 Scaled Pearson X2        399        401.0000          1.0050
                 Log Likelihood                     -266.8941


            Algorithm converged.
```

```
          replication of selected results in table 3: SVR<=>variables of interest

                                  The GENMOD Procedure

                             Analysis Of Parameter Estimates

                                     Standard     Wald 95% Confidence       Chi-
    Parameter    DF    Estimate        Error           Limits             Square    Pr > ChiSq

    Intercept     1     -0.7369       0.0882     -0.9099     -0.5640       69.78        <.0001
    male          1     -0.2950       0.1213     -0.5326     -0.0573        5.92        0.0150
    Scale         0      1.0000       0.0000      1.0000      1.0000

NOTE: The scale parameter was held fixed.


                            LR Statistics For Type 3 Analysis

                                                 Chi -
                         Source          DF     Square    Pr > ChiSq

                         male             1       5.64        0.0175


                                Contrast Estimate Results

                                      Standard                                          Chi-
 Label                 Estimate          Error      Alpha      Confidence Limits      Square    Pr > ChiSq

 Beta male              -0.2950         0.1213       0.05     -0.5326     -0.0573       5.92        0.0150
 Exp(Beta male)          0.7446         0.0903       0.05      0.5871      0.9443
```

```
       replication of selected results in table 3: SVR<=>variables of interest


                                The GENMOD Procedure

                                 Model Information

                        Data Set                WORK.ALL24_FU2
                        Distribution                  Binomial
                        Link Function                      Log
                        Dependent Variable                 svr


                      Number of Observations Read          401
                      Number of Observations Used          401
                      Number of Events                     160
                      Number of Trials                     401


                                  Response Profile

                             Ordered                Total
                               Value    svr     Frequency

                                   1    1             160
                                   2    0             241

PROC GENMOD is modeling the probability that svr='1'.


                               Parameter Information

                             Parameter       Effect

                             Prm1            Intercept
                             Prm2            age_5


                       Criteria For Assessing Goodness Of Fit

            Criterion                 DF           Value        Value/DF

            Deviance                 399        538.0465          1.3485
            Scaled Deviance          399        538.0465          1.3485
            Pearson Chi-Square       399        400.9357          1.0049
            Scaled Pearson X2        399        400.9357          1.0049
            Log Likelihood                     -269.0233


        Algorithm converged.
```

```
          replication of selected results in table 3: SVR<=>variables of interest


                                      The GENMOD Procedure

                                  Analysis Of Parameter Estimates

                                             Standard     Wald 95% Confidence       Chi-
     Parameter    DF    Estimate      Error           Limits           Square    Pr > ChiSq

     Intercept     1     -0.4962      0.3525     -1.1872      0.1947       1.98        0.1593
     age_5         1     -0.0441      0.0368     -0.1162      0.0280       1.44        0.2305
     Scale         0      1.0000      0.0000      1.0000      1.0000

OTE: The scale parameter was held fixed.


                                  LR Statistics For Type 3 Analysis

                                                  Chi -
                              Source           DF     Square    Pr > ChiSq

                              age_5             1       1.38        0.2396


                                     Contrast Estimate Results

                                       Standard                                                Chi-
abel                   Estimate          Error      Alpha       Confidence Limits      Square    Pr > ChiSq

eta age_5               -0.0441         0.0368       0.05      -0.1162      0.0280       1.44        0.2305
xp(Beta age_5)           0.9568         0.0352       0.05       0.8903      1.0284
```

```
           replication of selected results in table 3: SVR<=>variables of interest

                                 The GENMOD Procedure

                                  Model Information

                         Data Set                WORK.ALL24_FU2
                         Distribution                  Binomial
                         Link Function                      Log
                         Dependent Variable                 svr


                       Number of Observations Read         401
                       Number of Observations Used         395
                       Number of Events                    157
                       Number of Trials                    395
                       Missing Values                        6


                                  Response Profile

                              Ordered                Total
                                Value    svr    Frequency

                                    1    1            157
                                    2    0            238

PROC GENMOD is modeling the probability that svr='1'.


                                Parameter Information

                              Parameter       Effect

                              Prm1            Intercept
                              Prm2            hsorless


                         Criteria For Assessing Goodness Of Fit

              Criterion                 DF           Value        Value/DF

              Deviance                 393        524.2455          1.3340
              Scaled Deviance          393        524.2455          1.3340
              Pearson Chi-Square       393        395.0000          1.0051
              Scaled Pearson X2        393        395.0000          1.0051
              Log Likelihood                     -262.1228


          Algorithm converged.
```

```
            replication of selected results in table 3: SVR<=>variables of interest


                                  The GENMOD Procedure

                              Analysis Of Parameter Estimates

                                          Standard     Wald 95% Confidence       Chi-
    Parameter    DF    Estimate       Error           Limits             Square    Pr > ChiSq

    Intercept     1     -0.8004      0.0721     -0.9417     -0.6591     123.28        <.0001
    hsorless      1     -0.3367      0.1361     -0.6034     -0.0700       6.12        0.0133
    Scale         0      1.0000      0.0000      1.0000      1.0000

OTE: The scale parameter was held fixed.


                             LR Statistics For Type 3 Analysis

                                                Chi -
                           Source           DF     Square    Pr > ChiSq

                           hsorless          1       6.61        0.0101


                                 Contrast Estimate Results

                                      Standard                                         Chi -
 Label                    Estimate       Error    Alpha    Confidence Limits      Square   Pr > ChiSq

 Beta hsorless             -0.3367      0.1361     0.05    -0.6034    -0.0700       6.12        0.0133
 Exp(Beta hsorless)         0.7141      0.0972     0.05     0.5470     0.9324
```

```
             replication of selected results in table 3: SVR<=>variables of interest

                                       The GENMOD Procedure

                                        Model Information

                               Data Set                   WORK.ALL24_FU2
                               Distribution                     Binomial
                               Link Function                         Log
                               Dependent Variable                    svr


                             Number of Observations Read         401
                             Number of Observations Used         398
                             Number of Events                    158
                             Number of Trials                    398
                             Missing Values                        3


                                         Response Profile

                                    Ordered                Total
                                      Value    svr    Frequency

                                          1    1            158
                                          2    0            240

PROC GENMOD is modeling the probability that svr='1'.


                                      Parameter Information

                                    Parameter       Effect

                                    Prm1            Intercept
                                    Prm2            wgtkg_5


                              Criteria For Assessing Goodness Of Fit

                   Criterion                 DF           Value        Value/DF

                   Deviance                 396        521.3397          1.3165
                   Scaled Deviance          396        521.3397          1.3165
                   Pearson Chi-Square       396        398.2294          1.0056
                   Scaled Pearson X2        396        398.2294          1.0056
                   Log Likelihood                     -260.6699


           Algorithm converged.
```

```
          replication of selected results in table 3: SVR<=>variables of interest

                                   The GENMOD Procedure

                               Analysis Of Parameter Estimates

                                 Standard     Wald 95% Confidence       Chi-
    Parameter    DF    Estimate       Error           Limits            Square    Pr > ChiSq

    Intercept     1      0.1589      0.2919      -0.4131      0.7309       0.30        0.5862
    wgtkg_5       1     -0.0634      0.0177      -0.0980     -0.0287      12.87        0.0003
    Scale         0      1.0000      0.0000       1.0000      1.0000

NOTE: The scale parameter was held fixed.


                              LR Statistics For Type 3 Analysis

                                                   Chi -
                           Source           DF     Square    Pr > ChiSq

                           wgtkg_5           1      13.39        0.0003


                                 Contrast Estimate Results

                                   Standard                                     Chi-
  Label                 Estimate      Error    Alpha    Confidence Limits     Square   Pr > ChiSq

  Beta wgtkg_5           -0.0634     0.0177     0.05    -0.0980    -0.0287     12.87       0.0003
  Exp(Beta wgtkg_5)       0.9386     0.0166     0.05     0.9066     0.9717
```

```
         replication of selected results in table 3: SVR<=>variables of interest


                            The GENMOD Procedure

                             Model Information

                   Data Set                  WORK.ALL24_FU2
                   Distribution                    Binomial
                   Link Function                        Log
                   Dependent Variable                   svr


                 Number of Observations Read         401
                 Number of Observations Used         394
                 Number of Events                    155
                 Number of Trials                    394
                 Missing Values                        7


                              Response Profile

                       Ordered                Total
                         Value    svr    Frequency

                             1    1            155
                             2    0            239

ROC GENMOD is modeling the probability that svr='1'.


                           Parameter Information

                       Parameter       Effect

                       Prm1            Intercept
                       Prm2            bmi


                   Criteria For Assessing Goodness Of Fit

           Criterion                 DF           Value        Value/DF

           Deviance                 392        523.8940          1.3365
           Scaled Deviance          392        523.8940          1.3365
           Pearson Chi-Square       392        394.7328          1.0070
           Scaled Pearson X2        392        394.7328          1.0070
           Log Likelihood                     -261.9470


      Algorithm converged.
```

```
           replication of selected results in table 3: SVR<=>variables of interest


                                     The GENMOD Procedure

                                Analysis Of Parameter Estimates

                                          Standard     Wald 95% Confidence        Chi-
       Parameter    DF    Estimate        Error           Limits              Square    Pr > ChiSq

       Intercept     1     -0.2558       0.3353     -0.9131       0.4014        0.58        0.4455
       bmi           1     -0.0232       0.0116     -0.0460      -0.0005        4.00        0.0456
       Scale         0      1.0000       0.0000      1.0000       1.0000

NOTE: The scale parameter was held fixed.


                               LR Statistics For Type 3 Analysis

                                                       Chi -
                              Source           DF     Square    Pr > ChiSq

                              bmi               1       4.26        0.0390


                                  Contrast Estimate Results

                                       Standard                                           Chi-
 Label               Estimate             Error       Alpha      Confidence Limits      Square    Pr > ChiSq

 Beta bmi             -0.0232          0.0116        0.05     -0.0460     -0.0005        4.00        0.0456
 Exp(Beta bmi)         0.9770          0.0114        0.05      0.9550      0.9995
```

```
             replication of selected results in table 3: SVR<=>variables of interest

                                        The GENMOD Procedure

                                         Model Information

                               Data Set                WORK.ALL24_FU2
                               Distribution                  Binomial
                               Link Function                      Log
                               Dependent Variable                 svr


                             Number of Observations Read         401
                             Number of Observations Used         401
                             Number of Events                    160
                             Number of Trials                    401


                                          Response Profile

                                     Ordered               Total
                                       Value    svr    Frequency

                                           1    1            160
                                           2    0            241

PROC GENMOD is modeling the probability that svr='1'.


                                       Parameter Information

                                     Parameter       Effect

                                     Prm1            Intercept
                                     Prm2            MXDIAB


                              Criteria For Assessing Goodness Of Fit

                  Criterion                 DF           Value        Value/DF

                   Deviance                399        532.0851          1.3335
                   Scaled Deviance         399        532.0851          1.3335
                   Pearson Chi-Square      399        401.0000          1.0050
                   Scaled Pearson X2       399        401.0000          1.0050
                   Log Likelihood                    -266.0426


           Algorithm converged.
```

```
          replication of selected results in table 3: SVR<=>variables of interest

                                   The GENMOD Procedure

                               Analysis Of Parameter Estimates

                                       Standard     Wald 95% Confidence       Chi-
    Parameter    DF    Estimate       Error           Limits             Square    Pr > ChiSq

    Intercept     1     -0.8678      0.0618      -0.9888      -0.7467     197.30        <.0001
    MXDIAB        1     -0.7164      0.3212      -1.3459      -0.0868       4.97        0.0257
    Scale         0      1.0000      0.0000       1.0000       1.0000

NOTE: The scale parameter was held fixed.


                              LR Statistics For Type 3 Analysis

                                                   Chi -
                           Source           DF     Square    Pr > ChiSq

                           MXDIAB            1       7.34        0.0067


                                  Contrast Estimate Results

                                   Standard                                           Chi -
Label                  Estimate       Error     Alpha      Confidence Limits        Square    Pr >
ChiSq

Beta mxdiab             -0.7164      0.3212      0.05     -1.3459      -0.0868        4.97
0.0257
Exp(Beta mxdiab)         0.4885      0.1569      0.05      0.2603       0.9169
```

```
         replication of selected results in table 3: SVR<=>variables of interest


                                  The GENMOD Procedure

                                   Model Information

                          Data Set                    WORK.ALL24_FU2
                          Distribution                      Binomial
                          Link Function                          Log
                          Dependent Variable                     svr


                        Number of Observations Read          401
                        Number of Observations Used          401
                        Number of Events                     160
                        Number of Trials                     401


                                    Response Profile

                               Ordered               Total
                                 Value    svr    Frequency

                                     1    1            160
                                     2    0            241

ROC GENMOD is modeling the probability that svr='1'.


                                 Parameter Information

                               Parameter       Effect

                               Prm1            Intercept
                               Prm2            MXBP


                         Criteria For Assessing Goodness Of Fit

                Criterion                 DF           Value        Value/DF

                Deviance                 399        534.4976          1.3396
                Scaled Deviance          399        534.4976          1.3396
                Pearson Chi-Square       399        401.0000          1.0050
                Scaled Pearson X2        399        401.0000          1.0050
                Log Likelihood                     -267.2488


         Algorithm converged.
```

```
           replication of selected results in table 3: SVR<=>variables of interest

                                   The GENMOD Procedure

                              Analysis Of Parameter Estimates

                                     Standard     Wald 95% Confidence       Chi-
     Parameter    DF    Estimate        Error           Limits            Square    Pr > ChiSq

     Intercept     1     -0.8303       0.0689     -0.9653     -0.6954     145.45        <.0001
     MXBP          1     -0.3081       0.1460     -0.5943     -0.0219       4.45        0.0348
     Scale         0      1.0000       0.0000      1.0000      1.0000

NOTE: The scale parameter was held fixed.


                             LR Statistics For Type 3 Analysis

                                                 Chi -
                             Source          DF     Square    Pr > ChiSq

                             MXBP             1       4.93        0.0264


                                Contrast Estimate Results

                                Standard                                          Chi -
 Label                Estimate     Error      Alpha       Confidence Limits      Square    Pr > ChiSq

 Beta mxbp             -0.3081    0.1460       0.05     -0.5943     -0.0219       4.45        0.0348
 Exp(Beta mxbp)         0.7348    0.1073       0.05      0.5520      0.9783
```

```
             replication of selected results in table 3: SVR<=>variables of interest


                                       The GENMOD Procedure

                                        Model Information

                               Data Set               WORK.ALL24_FU2
                               Distribution                 Binomial
                               Link Function                     Log
                               Dependent Variable                svr


                             Number of Observations Read          401
                             Number of Observations Used          393
                             Number of Events                     157
                             Number of Trials                     393
                             Missing Values                         8


                                         Response Profile

                                    Ordered                Total
                                      Value    svr     Frequency

                                          1    1             157
                                          2    0             236

PROC GENMOD is modeling the probability that svr='1'.


                                      Parameter Information

                                    Parameter       Effect

                                    Prm1            Intercept
                                    Prm2            currentsmk


                              Criteria For Assessing Goodness Of Fit

                   Criterion                 DF            Value        Value/DF

                   Deviance                 391         527.7539          1.3498
                   Scaled Deviance          391         527.7539          1.3498
                   Pearson Chi-Square       391         392.9999          1.0051
                   Scaled Pearson X2        391         392.9999          1.0051
                   Log Likelihood                      -263.8770


           Algorithm converged.
```

```
           replication of selected results in table 3: SVR<=>variables of interest


                                    The GENMOD Procedure

                                Analysis Of Parameter Estimates

                                     Standard     Wald 95% Confidence       Chi-
   Parameter      DF     Estimate       Error           Limits            Square    Pr > ChiSq

   Intercept       1      -0.9687      0.0817      -1.1287      -0.8086     140.65        <.0001
   currentsmk      1       0.1303      0.1247      -0.1140       0.3747       1.09        0.2958
   Scale           0       1.0000      0.0000       1.0000       1.0000

OTE: The scale parameter was held fixed.


                               LR Statistics For Type 3 Analysis

                                                       Chi-
                            Source            DF     Square    Pr > ChiSq

                            currentsmk         1       1.07        0.3008


                                  Contrast Estimate Results

                                      Standard                                        Chi-
Label                     Estimate       Error     Alpha    Confidence Limits     Square   Pr > ChiSq

Beta currentsmk             0.1303      0.1247      0.05    -0.1140     0.3747      1.09       0.2958
Exp(Beta currentsmk)        1.1392      0.1420      0.05     0.8923     1.4545
```

```
           replication of selected results in table 3: SVR<=>variables of interest

                                   The GENMOD Procedure

                                    Model Information

                          Data Set                 WORK.ALL24_FU2
                          Distribution                   Binomial
                          Link Function                       Log
                          Dependent Variable                  svr


                        Number of Observations Read         401
                        Number of Observations Used         399
                        Number of Events                    160
                        Number of Trials                    399
                        Missing Values                        2


                                     Response Profile

                               Ordered                Total
                                 Value    svr    Frequency

                                     1    1            160
                                     2    0            239

PROC GENMOD is modeling the probability that svr='1'.


                                  Parameter Information

                               Parameter       Effect

                               Prm1            Intercept
                               Prm2            alt_100


                          Criteria For Assessing Goodness Of Fit

                Criterion                 DF           Value        Value/DF

                Deviance                 397        536.9490          1.3525
                Scaled Deviance          397        536.9490          1.3525
                Pearson Chi-Square       397        399.0127          1.0051
                Scaled Pearson X2        397        399.0127          1.0051
                Log Likelihood                     -268.4745


          Algorithm converged.
```

```
          replication of selected results in table 3: SVR<=>variables of interest

                                   The GENMOD Procedure

                              Analysis Of Parameter Estimates

                                    Standard     Wald 95% Confidence       Chi-
     Parameter    DF    Estimate       Error           Limits            Square    Pr > ChiSq

     Intercept     1     -0.9614      0.0942     -1.1461     -0.7768     104.13        <.0001
     alt_100       1      0.0522      0.0753     -0.0954      0.1998       0.48        0.4881
     Scale         0      1.0000      0.0000      1.0000      1.0000

NOTE: The scale parameter was held fixed.


                             LR Statistics For Type 3 Analysis

                                                 Chi -
                           Source           DF     Square    Pr > ChiSq

                           alt_100           1       0.44        0.5086


                                 Contrast Estimate Results

                                  Standard                                      Chi-
   Label                 Estimate      Error    Alpha    Confidence Limits    Square   Pr > ChiSq

   Beta alt_100            0.0522     0.0753     0.05    -0.0954     0.1998     0.48       0.4881
   Exp(Beta alt_100)       1.0536     0.0793     0.05     0.9090     1.2211
```

```
          replication of selected results in table 3: SVR<=>variables of interest


                                The GENMOD Procedure

                                 Model Information

                        Data Set                WORK.ALL24_FU2
                        Distribution                  Binomial
                        Link Function                      Log
                        Dependent Variable                 svr


                      Number of Observations Read         401
                      Number of Observations Used         399
                      Number of Events                    160
                      Number of Trials                    399
                      Missing Values                        2


                                  Response Profile

                             Ordered                Total
                               Value    svr    Frequency

                                   1    1            160
                                   2    0            239

PROC GENMOD is modeling the probability that svr='1'.


                               Parameter Information

                             Parameter       Effect

                             Prm1            Intercept
                             Prm2            ast_100


                        Criteria For Assessing Goodness Of Fit

               Criterion                 DF           Value        Value/DF

               Deviance                 397        533.3470          1.3434
               Scaled Deviance          397        533.3470          1.3434
               Pearson Chi-Square       397        399.0699          1.0052
               Scaled Pearson X2        397        399.0699          1.0052
               Log Likelihood                     -266.6735


        Algorithm converged.
```

```
          replication of selected results in table 3: SVR<=>variables of interest

                                   The GENMOD Procedure

                             Analysis Of Parameter Estimates

                                     Standard     Wald 95% Confidence       Chi-
    Parameter    DF    Estimate        Error           Limits             Square    Pr > ChiSq

    Intercept     1     -0.7342       0.1099     -0.9496     -0.5188       44.64        <.0001
    ast_100       1     -0.2853       0.1588     -0.5966      0.0259        3.23        0.0724
    Scale         0      1.0000       0.0000      1.0000      1.0000

NOTE: The scale parameter was held fixed.


                           LR Statistics For Type 3 Analysis

                                                  Chi -
                             Source          DF     Square    Pr > ChiSq

                             ast_100          1       4.04        0.0445


                                 Contrast Estimate Results

                                   Standard                                       Chi -
   Label                Estimate      Error     Alpha    Confidence Limits      Square   Pr > ChiSq

   Beta ast_100          -0.2853     0.1588      0.05    -0.5966      0.0259      3.23       0.0724
   Exp(Beta ast_100)      0.7518     0.1194      0.05     0.5507      1.0263
```

```
             replication of selected results in table 3: SVR<=>variables of interest

                                       The GENMOD Procedure

                                        Model Information

                               Data Set                  WORK.ALL24_FU2
                               Distribution                    Binomial
                               Link Function                        Log
                               Dependent Variable                   svr


                             Number of Observations Read         401
                             Number of Observations Used         395
                             Number of Events                    158
                             Number of Trials                    395
                             Missing Values                        6


                                         Response Profile

                                    Ordered              Total
                                      Value    svr    Frequency

                                          1    1            158
                                          2    0            237

PROC GENMOD is modeling the probability that svr='1'.


                                      Parameter Information

                                    Parameter       Effect

                                    Prm1            Intercept
                                    Prm2            HGB


                              Criteria For Assessing Goodness Of Fit

                   Criterion                 DF           Value        Value/DF

                   Deviance                 393        530.5994          1.3501
                   Scaled Deviance          393        530.5994          1.3501
                   Pearson Chi-Square       393        395.1132          1.0054
                   Scaled Pearson X2        393        395.1132          1.0054
                   Log Likelihood                     -265.2997


           Algorithm converged.
```

```
          replication of selected results in table 3: SVR<=>variables of interest


                                  The GENMOD Procedure

                              Analysis Of Parameter Estimates

                                      Standard     Wald 95% Confidence       Chi-
     Parameter    DF     Estimate        Error           Limits            Square    Pr > ChiSq

     Intercept     1      -0.1984       0.6885     -1.5479      1.1511       0.08        0.7732
     HGB           1      -0.0492       0.0473     -0.1419      0.0435       1.08        0.2984
     Scale         0       1.0000       0.0000      1.0000      1.0000

NOTE: The scale parameter was held fixed.


                             LR Statistics For Type 3 Analysis

                                                     Chi -
                             Source           DF     Square    Pr > ChiSq

                             HGB               1       1.08        0.2987


                                  Contrast Estimate Results

                                    Standard                                       Chi -
 Label               Estimate          Error      Alpha       Confidence Limits     Square    Pr > ChiSq

 Beta hgb             -0.0492         0.0473       0.05     -0.1419       0.0435      1.08        0.2984
 Exp(Beta hgb)         0.9520         0.0450       0.05      0.8677       1.0445
```

```
             replication of selected results in table 3: SVR<=>variables of interest

                                       The GENMOD Procedure

                                        Model Information

                               Data Set                WORK.ALL24_FU2
                               Distribution                  Binomial
                               Link Function                      Log
                               Dependent Variable                 svr


                             Number of Observations Read         401
                             Number of Observations Used         395
                             Number of Events                    158
                             Number of Trials                    395
                             Missing Values                        6


                                         Response Profile

                                    Ordered               Total
                                      Value    svr    Frequency

                                          1    1            158
                                          2    0            237

PROC GENMOD is modeling the probability that svr='1'.


                                      Parameter Information

                                    Parameter       Effect

                                    Prm1            Intercept
                                    Prm2            WBC


                              Criteria For Assessing Goodness Of Fit

                   Criterion                 DF           Value        Value/DF

                   Deviance                 393        529.0558          1.3462
                   Scaled Deviance          393        529.0558          1.3462
                   Pearson Chi-Square       393        394.5044          1.0038
                   Scaled Pearson X2        393        394.5044          1.0038
                   Log Likelihood                     -264.5279


           Algorithm converged.
```

```
              replication of selected results in table 3: SVR<=>variables of interest

                                          The GENMOD Procedure

                                     Analysis Of Parameter Estimates

                                           Standard     Wald 95% Confidence       Chi-
    Parameter    DF    Estimate       Error           Limits             Square    Pr > ChiSq

    Intercept     1     -1.1977      0.1840     -1.5583     -0.8370       42.36        <.0001
    WBC           1      0.0441      0.0260     -0.0069      0.0951        2.87        0.0903
    Scale         0      1.0000      0.0000      1.0000      1.0000

NOTE: The scale parameter was held fixed.


                                    LR Statistics For Type 3 Analysis

                                                         Chi-
                             Source           DF      Square    Pr > ChiSq

                             WBC               1        2.62        0.1053


                                      Contrast Estimate Results

                                        Standard                                        Chi-
Label               Estimate       Error      Alpha       Confidence Limits      Square    Pr > ChiSq

Beta wbc              0.0441      0.0260       0.05     -0.0069      0.0951        2.87        0.0903
Exp(Beta wbc)         1.0451      0.0272       0.05      0.9931      1.0997
```

```
             replication of selected results in table 3: SVR<=>variables of interest

                                       The GENMOD Procedure

                                        Model Information

                               Data Set             WORK.ALL24_FU2
                               Distribution               Binomial
                               Link Function                   Log
                               Dependent Variable              svr


                             Number of Observations Read         401
                             Number of Observations Used         399
                             Number of Events                    160
                             Number of Trials                    399
                             Missing Values                        2


                                         Response Profile

                                    Ordered                Total
                                      Value    svr    Frequency

                                          1    1            160
                                          2    0            239

PROC GENMOD is modeling the probability that svr='1'.


                                      Parameter Information

                                   Parameter       Effect

                                   Prm1            Intercept
                                   Prm2            blog10vload


                              Criteria For Assessing Goodness Of Fit

                    Criterion                 DF           Value        Value/DF

                    Deviance                 397        524.7746          1.3219
                    Scaled Deviance          397        524.7746          1.3219
                    Pearson Chi-Square       397        398.7280          1.0044
                    Scaled Pearson X2        397        398.7280          1.0044
                    Log Likelihood                     -262.3873


           Algorithm converged.
```

```
           replication of selected results in table 3: SVR<=>variables of interest

                                    The GENMOD Procedure

                               Analysis Of Parameter Estimates

                                        Standard     Wald 95% Confidence        Chi-
   Parameter       DF    Estimate        Error           Limits              Square    Pr > ChiSq

   Intercept        1      0.6195       0.2814       0.0681       1.1709       4.85        0.0277
   blog10vload      1     -0.2476       0.0480      -0.3416      -0.1536      26.64        <.0001
   Scale            0      1.0000       0.0000       1.0000       1.0000

NOTE: The scale parameter was held fixed.


                              LR  Statistics For Type 3 Analysis

                                                          Chi -
                         Source             DF     Square    Pr > ChiSq

                         blog10vload         1      12.61        0.0004


                                  Contrast Estimate Results

                                      Standard                                          Chi -
Label                       Estimate       Error    Alpha    Confidence Limits     Square   Pr > ChiSq

Beta blog10vload             -0.2476      0.0480     0.05    -0.3416    -0.1536    26.64        <.0001
Exp(Beta blog10vload)         0.7807      0.0374     0.05     0.7106     0.8576
```