Appendix 1: Technical details

Let patients be indexed from $i = 1$ to 761 and x index the treatment arms. T_i^s is the time from randomization to first viral suppression, T_i^r is the time from randomization to viral rebound, and T_i^c is the time from randomization to censoring at loss to follow-up, day 1012, or death. Define $R_i^s = \min(T_i^s, T_i^c)$ and $R_i^r = \min(T_i^r, T_i^c)$. δ_i^s is an indicator that $T_i^s < T_i^c$ and δ_i^r is an indicator that $T_i^r < T_i^c$. n_x is the number of patients in arm x. For each arm, we estimate the time to initial suppression using the nonparametric Kaplan-Meier product limit estimator of the survivor function: $\hat{S}_x^s(t) = \prod_{t_{x,j}<t}(1 - d_{x,j}^s / y_{x,j}^s)$, where j indexes the distinct suppression event times, $d_{x,j}^s$ is the number of suppression events at time $t_{x,j}$, and $y_{x,j}^s$ is the number of patients at risk of initial suppression at time $t_{x,j}$. Similarly, we estimate the time from randomization to viral rebound as $\hat{S}_x^r(t) = \prod_{t_{x,j}<t}(1 - d_{x,j}^r / y_{x,j}^r)$, where j indexes the distinct rebound event times, $d_{x,j}^r$ is the number of viral rebound events at time $t_{x,j}$ and $y_{x,j}^r$ is the number of patients at risk for viral rebound at time $t_{x,j}$.

The probability that a patient with exposure x has a suppressed viral load at time t is the difference of these two curves: $\hat{G}_x(t) = \hat{S}_x^r(t) - \hat{S}_x^s(t)$. Rather than choose a time point at which to estimate this probability, we can summarize this probability over follow-up interval $[0, \tau]$ as the mean time spent in a state of suppression, estimated as $A_x = \int_0^\tau \hat{G}_x(u)du$. The mean time in a state of suppression can be compared between exposure groups as $DS = \int_0^\tau \{\hat{G}_1(u) - \hat{G}_0(u)\}\mathrm{d}u$, which can be calculated as the Riemann sum $DS = \sum_{k=1}^{\tau} \{\hat{G}_1(u_k) - \hat{G}_0(u_k)\} \Delta u_k$.

The variance of DS can be estimated as

$$\widehat{Var}(DS) = \hat{V}_0 + \hat{V}_1,$$

where

$$\hat{V}_x = \frac{1}{n_x^2} \sum_{i:\, X_i = x} \left[\int_0^{\tau} P_i(u) du \right]^2 .$$

$$P_i(t) = n_x \hat{S}_x^s(t) \left\{ \int_0^t \frac{1}{y_x^s(u)} \delta_i^s \, du - \int_0^t \frac{Y_i^s(u)}{\left(y_x^s(u)\right)^2} d_u^s du \right\}$$

$$- n_x \hat{S}_x^r(t) \left\{ \int_0^t \frac{1}{y_x^r(u)} \delta_i^r \, du - \int_0^t \frac{Y_i^r(u)}{\left(y_x^r(u)\right)^2} d_u^r du \right\},$$

Where $Y_i^s(u)$ is an indicator for individual i of being at risk of suppression at time u, $y_x^s(u) = \sum_{i:X_i=x} Y_i^s(u)$ and $Y_i^r(u)$ is an indicator of being at risk of rebound at time u, $y_x^r(u) = \sum_{i:X_i=x} Y_i^r(u)$.

Appendix 2: SAS code

```
*****************************************************************************
Code for viral suppression manuscript illustrating approach outlined by
Gouskova 2014

start with dataset d with the following variables:
t1: time from randomization to suppression
d1: indicator of suppression at time t1
t2: time from randomization to rebound
d2: indicator of rebound at t2
x: exposure indicator
****************************************************************************;

*Estimate S(Suppression);
proc phreg data=d noprint; model t1*d1(0)=; strata x; baseline out=sup(keep=
x t1 s1) survival=s1;
data sup; set sup; t=t1;

*Estimate S(Rebound);
proc phreg data=d noprint; model t2*d2(0)=; strata x; baseline out=reb(keep=
x t2 s2) survival=s2;
data reb; set reb; t=t2;

*Make G(t) = S(Reb,t)-S(Sup,t);
proc sort data=sup; by x t;
proc sort data=reb; by x t;
data times;  do x=0 to 1; do t=0 to 1012; output; end;  end;
data g;
      merge sup reb times;
      by  x t;
      keep  x t s1 s2;
data g; set g; by  x t; retain holdr1 holdr2;
      if first.x then do; holdr1=0; holdr2=0; end;
      if s1=. then s1=holdr1; if s2=. then s2=holdr2;
      delta=s2-s1;
      output;
      holdr1=s1; holdr2=s2;
run;

data three; set g; if x=0; delta0=delta; keep  t delta0;
data four; set g; if x=1; delta1=delta; keep  t delta1;
data deltag; merge three four; by  t; retain sum0 sum1 0;
      if first.b then do; sum0=0; sum1=0; end;
      sum0=sum0+delta0; sum1=sum1+delta1;
      if t=0 then do; g0=0; g1=0; end;
      if t>0 then do; g0=(sum0/t)*1012; g1=(sum1/t)*1012;end;
      delta=g1-g0;
      if t=1012 then output;
proc print data=deltag noobs;  var   g0  g1 delta;
      title "Mean time in suppression and difference for x=0 and =1";
run;

*Variance;
proc phreg data=d  atrisk noprint; model t1*d1(0)=; strata x;
```

```
output out=ars(keep=x t1 ys) num_left=ys ;
baseline out=sup(keep=x t1 s) survival=s; run;
proc sort data=ars; by x t1;
data sup; set sup; by x t1; if last.t1;* get rid of duplicate record at t=0;
data ars; set ars; by x t1; if first.t1;
data shell; do x=0 to 1; do t1=1 to 1012;t2=t1; output; end; end;run;

data sup2; merge sup ars; by x t1;
retain s2; if s ne . then s2=s; else s2=s2; drop s;
rename s2=s; run;

data sup2;
      set sup2;
      by x t1;
      retain chm1 0 sm1 1 tm1 0 nsu;
      if first.x then  do; nsu=0; chm1=0; sm1=1; end;
      if s>0 then ch=-log(s);
      nsu=round((ch-chm1)*ys,1);*above recreates y because SAS does not
output;
      sst=s;
      t=t1;
      output;
      chm1=ch; sm1=s;
      keep x t nsu ys  sst ;
run;

proc phreg data=d  atrisk noprint; model t2*d2(0)=; strata x;
output out=arr(keep=x t2 yr) num_left=yr ;
baseline out=reb(keep=x t2 s) survival=s; run;
proc sort data=arr; by x t2;
data reb; set reb; by x t2; if last.t2;* get rid of duplicate record at t=0;
data arr; set arr; by x t1; if first.t2;
data shell; do x=0 to 1; do t1=1 to 1012;t2=t1; t=t1; output; end; end;run;

data reb2; merge reb arr; by x t2;
retain s2; if s ne . then s2=s; else s2=s2; drop s;
rename s2=s; run;

data reb2;
      set reb2;
      by x t2;
      retain chm1 0 sm1 1 tm1 0 nru;
      if first.x then  do; nru=0; chm1=0; sm1=1;  end;
      if s>0 then ch=-log(s);
      nru=round((ch-chm1)*yr,1);* recreates nru because SAS does not output;
      ssr=s;
      t=t2;
      output;
      chm1=ch; sm1=s;
      keep x t ssr nru yr  ;
run;



data ltp4; merge reb2 sup2 shell; by x t;
retain yr2 nru2 ys2 nsu2 ssr2 sst2;
      if yr ne . then yr2=yr; else  yr2=yr2;
```

```
      if nru ne . then nru2=nru; else nru2=0;
      if ys ne . then ys2=ys; else ys2=ys2;
      if nsu ne . then nsu2=nsu; else nsu2=0;
      if sst ne . then sst2=sst; else sst2=sst2;
      if ssr ne . then ssr2=ssr; else ssr2=ssr2;
      if last.t then output;
drop yr nru ys nsu sst ssr t1 t2;
rename yr2=yr nru2=nru ys2=ys nsu2=nsu ssr2=ssr sst2=sst;
run;

data e; set d; do t=0 to 1012; output; end; run;
proc sort data=e; by x t;

data lte; merge e(in=ine)  ltp4 ; by x t ; if ine; run;
proc sort data=lte; by x id t; run;

data lte; set lte;
if t<=t1 then  ysui=1; else ysui=0;
if t<=t2 then yrui=1; else yrui=0;
if t1=t and d1=1 then nsi=1; else nsi=0;
if t2=t and d2=1 then nri=1; else nri=0;
if t<=1012 then output;
run;

*count number in each arm;
proc sql; create table counts as select x, count(distinct(id)) as nx from lte
group by x;quit;
data lte2; merge  counts lte; by x; run;
proc sort data=lte2; by x id; run;

*get components of Pi(t);
data lte2; set lte2; by x id;
q1u=(1/ys)*nsi;
q2u=(ysui/ys**2)*nsu;
q3u=(1/yr)*nri;
q4u=(yrui/yr**2)*nru;
run;

*sum components of Pi(t) over time (until time t) and calculate Pi(t);
data lte3; set lte2; by x id;
retain q1t q2t q3t q4t;
if first.id then do; q1t=0; q2t=0; q3t=0; q4t=0; end;
q1t=q1t+q1u; q2t=q2t+q2u; q3t=q3t+q3u; q4t=q4t+q4u;
pit=nx*sst*(q1t-q2t)-(nx*ssr*(q3t-q4t));
run;

data lte4; set lte3; by x id;
retain q6i;
if first.id then q6i=0;
q6i=pit+q6i;
if t>0 then p=(q6i);
if last.id then output;
run;
data lte5; set lte4; by x;
retain q8; if first.x then q8=0;
q7=p**2;
q8=q7+q8;
```

```
if last.x then output;
run;
data lte6; set lte5 ; by x; retain v0 n0;
if x=0 then do; v0=(1/nx**2)*q8; n0=nx;end;
if x=1 then do; v1=(1/nx**2)*q8; n1=nx;end;
if x=1 then output;
keep v0 v1 n0 n1;
run;
data lte7; set lte6;
se1=sqrt(v1); se0=sqrt(v0);
sediff=sqrt(v0+v1);
run;
ods listing;

*get estimate and 95% CI;
data est; merge reb2  sup2 shell ;by x t;
data est; set est;by x; retain sr st;if first.x then do; sr=1; st=1; end;
if ssr ne . then sr=ssr;
if sst ne . then st=sst;
run;

data estx1; set est;  if x=1 ; gt1=sr-st; run;
data estx0; set est;  if x=0 ; gt0=sr-st; run;
data estx; merge estx1 estx0 end=eof; by t;
if last.t;
 retain sum0 0 sum1 0 g0 g1;
      sum0=sum(gt0,sum0);
      sum1=sum(gt1,sum1);
      if t=0 then do; g0=0; g1=0; end;
      else do;
            g0=sum0/(t)*1012;
            g1=sum1/(t)*1012;
      end;
      if eof then output;
run;
data estx; merge estx(keep=g0 g1) lte7(keep=sediff se1 se0);
      gtdiff=g1-g0;
      lcl=gtdiff-1.96*sediff;
      ucl=gtdiff+1.96*sediff;
proc print data=estx; var g1 g0 gtdiff lcl ucl;
title "Mean days suppressed for 3- and 4-drug arms with different and 95%
CI";
run;
```