The Effectiveness of Combinatorial Antiretroviral Treatments on Treating Patients with HIV-1
STAT 678 Final Project

Kyle Lee   Yunhao Sun

Department of Mathematics and Statistics
San Diego State University

August 11, 2018

Outline

Background

- Human immunodeficiency virus (HIV) destroys CD4 cells. There are 5 stages of diagnosis (0,1,2,3, unknown) and AIDS is Stage 3.

- Most likely thought to originate from a type of chimpanzee in Central Africa as far as late 1800s[1].

- Before treatment options existed in the mid 1990s, HIV-inflicted patients can live only for several years.

- Currently, no cure exists. But treatment options exist that minimize the viral counts.

- Treatment regimens proposed in mid 1990s and followed by more research. This dataset came from one of the first studies on AIDS treatments (1997).

- Study terminated early due to "demonstration ... of clinical superiority"[2] of the three-drug regimen over the two-drug regimen.

Background
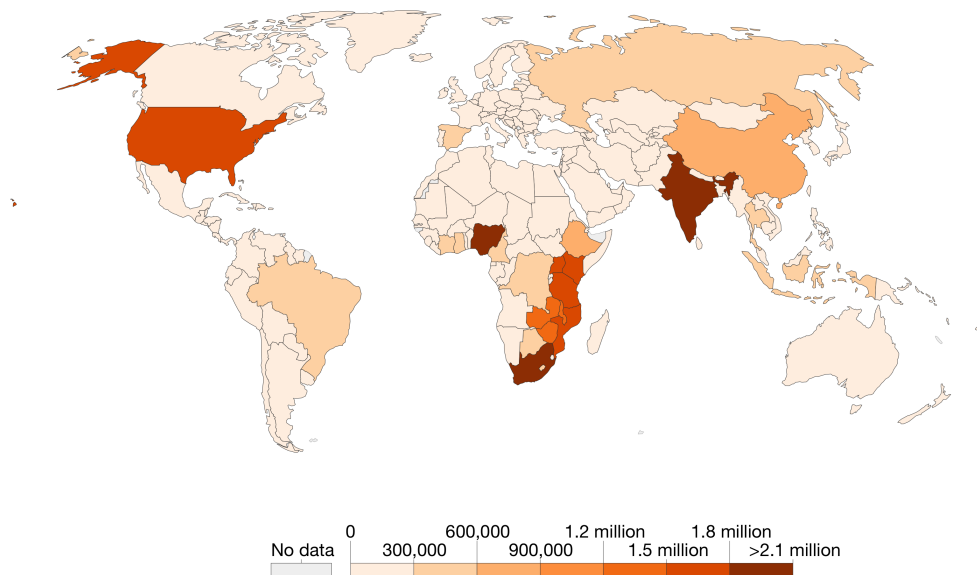
Figure 1: Source: *Our World in Data*
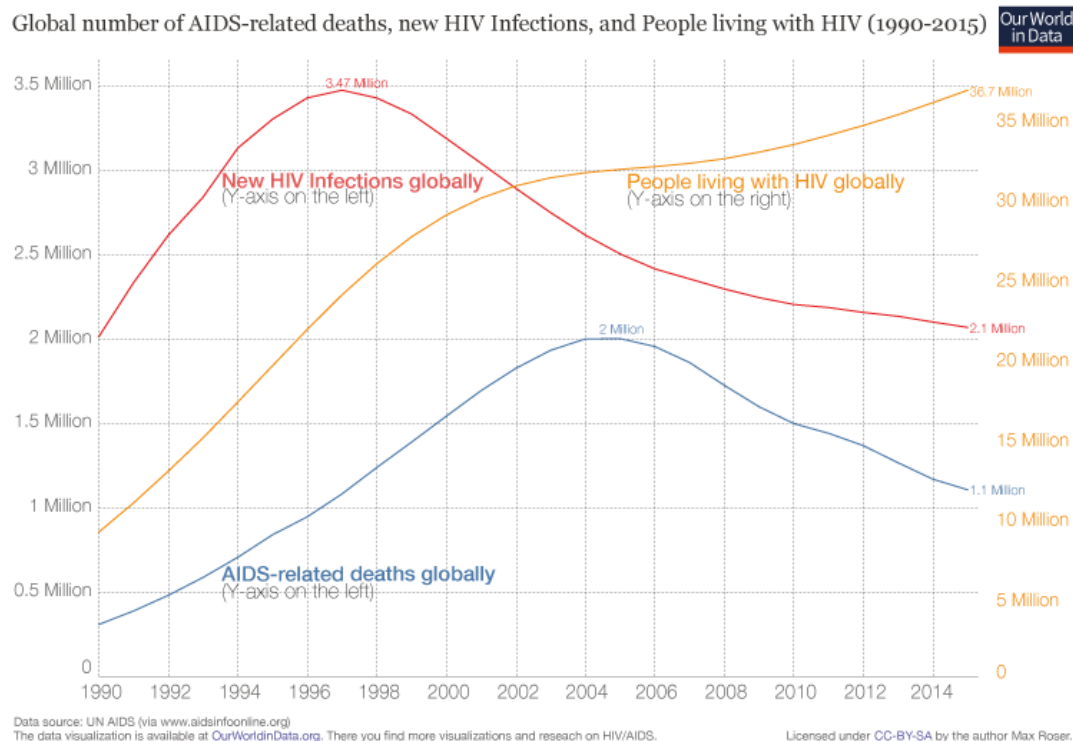
Background

Figure 2: Source: *Our World in Data*

Data

- Double-blind clinical trial with **two treatment regimes**:
  1. indinavir (IDV), zidovudine (ZDV) or stavudine (d4T) and lamivudine (3TC)
  2. zidovudine (ZDV) or stavudine (d4T) and lamivudine (3TC)
- **Mechanism of action:**
  Inhibit HIV's reverse transcriptase, the enzyme that the virus uses to make a DNA copy of its RNA.
- **Total observations:** 1151, 5 missing from original dataset
- Randomization stratified according to CD4 cell count at the time of the screening
- Variables: id, $time$, $censor$, $time_d$, $censor_d$, tx, $txgrp$, $strat2$, sex, $raceth$, $ivdrug$, $hemophil$, $karnof$, $cd4$, $priorzdv$, age

Variables

Name	Description	Codes/Values
id	Identification Code	1-1151
$time$	Time to AIDS diagnosis or death	Days
$censor$	Event indicator for AIDS/death	1/0
tx	Treatment indicator	1 (includes IDV)/0
$txgrp$	Treatment group indicator	1-4
$strat2$	CD4 stratum at screening	1/0
sex	Sex	1/0
$raceth$	Race/Ethnicity	1-6
$ivdrug$	IV drug use history	1-3
$hemophil$	Hemophiliac	1/0
$karnof$	Karnofsky Performance Scale	70, 80, 90, 100
cd	Baseline CD4 count	Cells/milliliter
$priorzdv$	Months of prior ZDV use	Months
age	Age at Enrollment	Years

Note: We do not use $time_d$ and $censor_d$ because most survive. Yay!

Summary Statistics

- 439(38.1%) patients had 50 or fewer CD4 cells per cubic millimeter and 712(61.9%) patients had 51 to 200 CD4 cells per cubic millimeter.

- 951(82.6%) of patients were male and 200(17.4%) were female.

- 596(51.8%) of patients were white non-hispanic, 327(28.4%) were black non-hispanic, 203(17.6%) were hispanic, 14(1.2%) were Asian or Pacific Islander, and 11(1.0%) were American Indian or Alaskan Native.

- Karnofsky Score: 32(2.8%) scored a 70, 182(15.8%) scored an 80, 541(47.0%) scored a 90, and 396(34.4%) scored a 100.
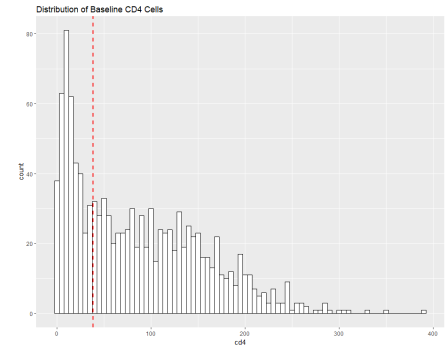
Summary Statistics
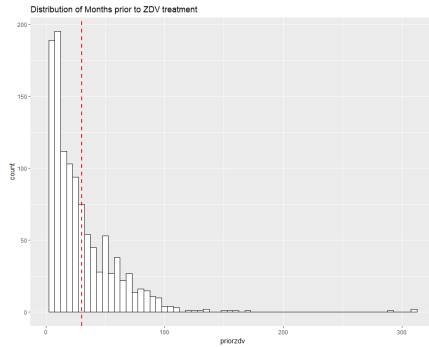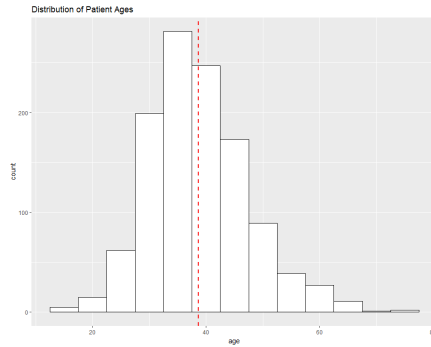
(a) Time Prior to ZDV   (b) Patient Ages   (c) Baseline CD4

Figure 3: TOTAL

Method

- Backward Variable Selection
    1. Fit a Cox model with all appropriate variables
    2. Find the largest p-value and remove that variable. Consider factored variables and their p-values.
    3. Repeat Steps 1 and 2 until variables are relatively significant

- Identifying functional form of covariates

- Assess the Proportional Hazards Assumption

- Identify Outliers

- Create final model

- Discussion

Backward Variable Selection

Figure 4: P-values for Model Selection

	Dependent variable: time					
	(1)	(2)	(3)	(4)	(5)	(6)
tx	$p = 0.003$***	$p = 0.003$***	$p = 0.003$***	$p = 0.003$***	$p = 0.003$***	$p = 0.002$***
strat2	$p = 0.857$	$p = 0.858$				
sex	$p = 0.490$	$p = 0.495$	$p = 0.494$	$p = 0.498$		
factor(raceth)2	$p = 0.190$	$p = 0.189$	$p = 0.191$	$p = 0.198$	$p = 0.231$	
factor(raceth)3	$p = 0.721$	$p = 0.722$	$p = 0.715$	$p = 0.705$	$p = 0.673$	
factor(raceth)4	$p = 0.173$	$p = 0.174$	$p = 0.173$	$p = 0.174$	$p = 0.169$	
factor(raceth)5	$p = 0.810$	$p = 0.789$	$p = 0.790$	$p = 0.789$	$p = 0.804$	
factor(ivdrug)2	$p = 0.463$	$p = 0.464$	$p = 0.476$	$p = 0.487$	$p = 0.489$	$p = 0.465$
factor(ivdrug)3	$p = 0.081$*	$p = 0.080$*	$p = 0.081$*	$p = 0.081$*	$p = 0.083$*	$p = 0.068$*
hemophil	$p = 0.873$					
karnof	$p = 0.00001$***	$p = 0.00001$***	$p = 0.00001$***	$p = 0.00001$***	$p = 0.00001$***	$p = 0.00001$***
cd4	$p = 0.0003$***	$p = 0.0002$***	$p = 0.000$***	$p = 0.000$***	$p = 0.000$***	$p = 0.000$***
priorzdv	$p = 0.774$	$p = 0.789$	$p = 0.797$			
age	$p = 0.040$**	$p = 0.041$**	$p = 0.042$**	$p = 0.043$**	$p = 0.050$**	$p = 0.060$*

Finding functional forms of covariates

Figure 5: Finding functional forms of covariates[3]
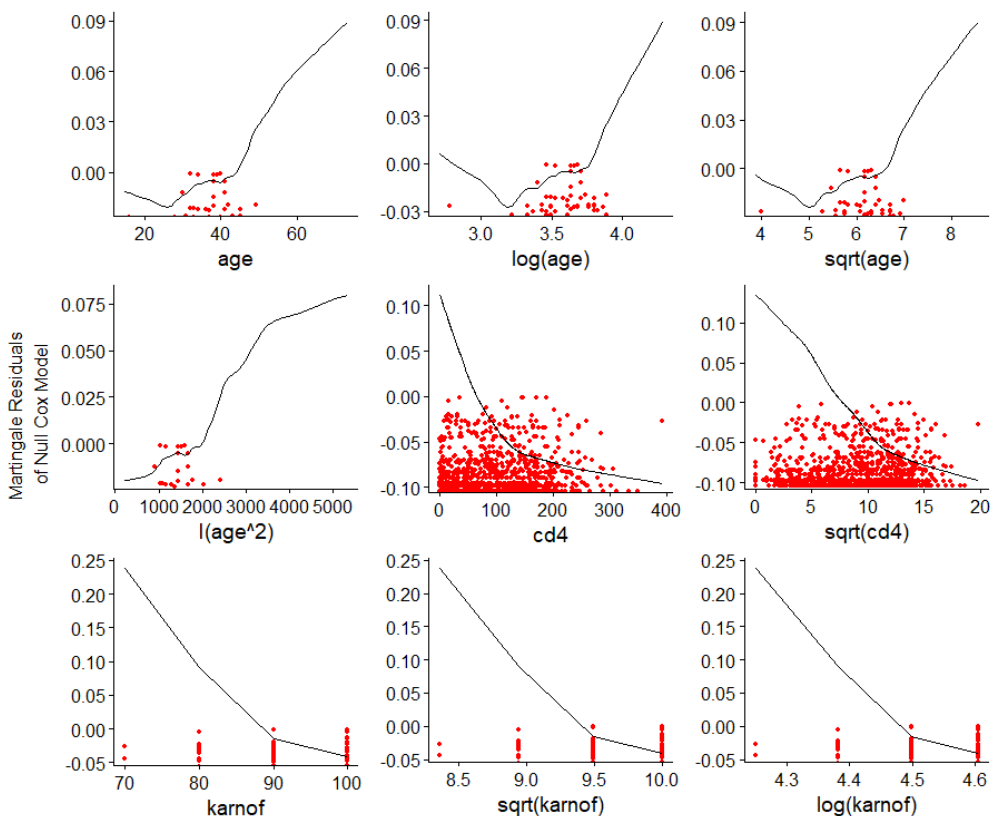
Testing Proportional Hazards Assumption
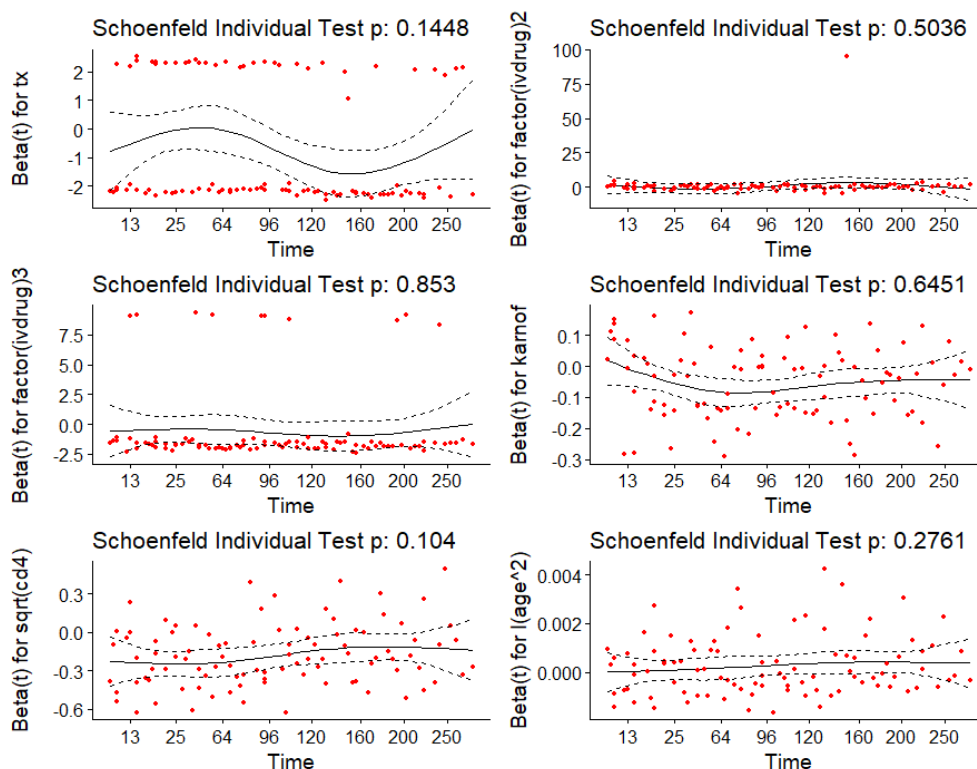
Figure 6: Testing PH Assumption Part 1

Figure 7: Testing PH Assumption Part 2
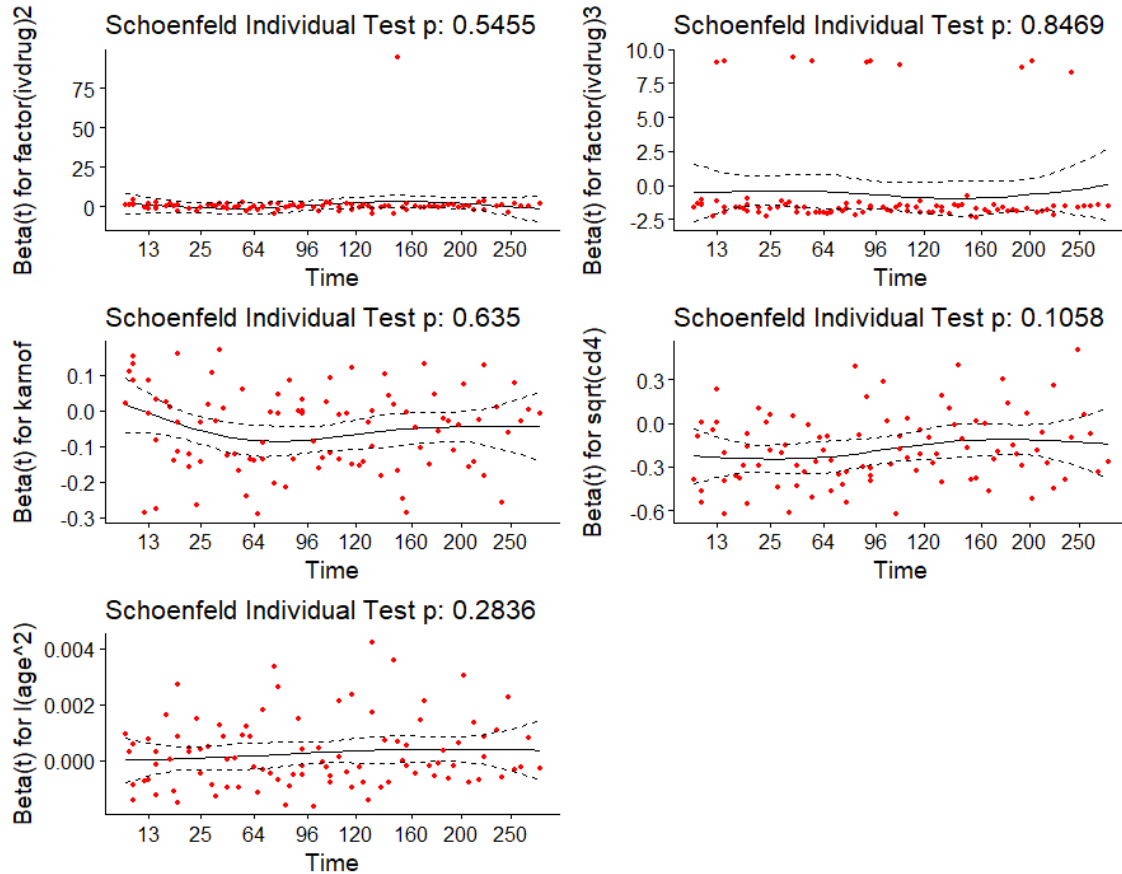
Outlier Analysis

Score Residuals of factor(ivdrug)2

Figure 9: Score Residuals for factor(ivdrug)3

Score Residuals of factor(ivdrug)3

Figure 10: Score Residuals for Karnofsky Scores
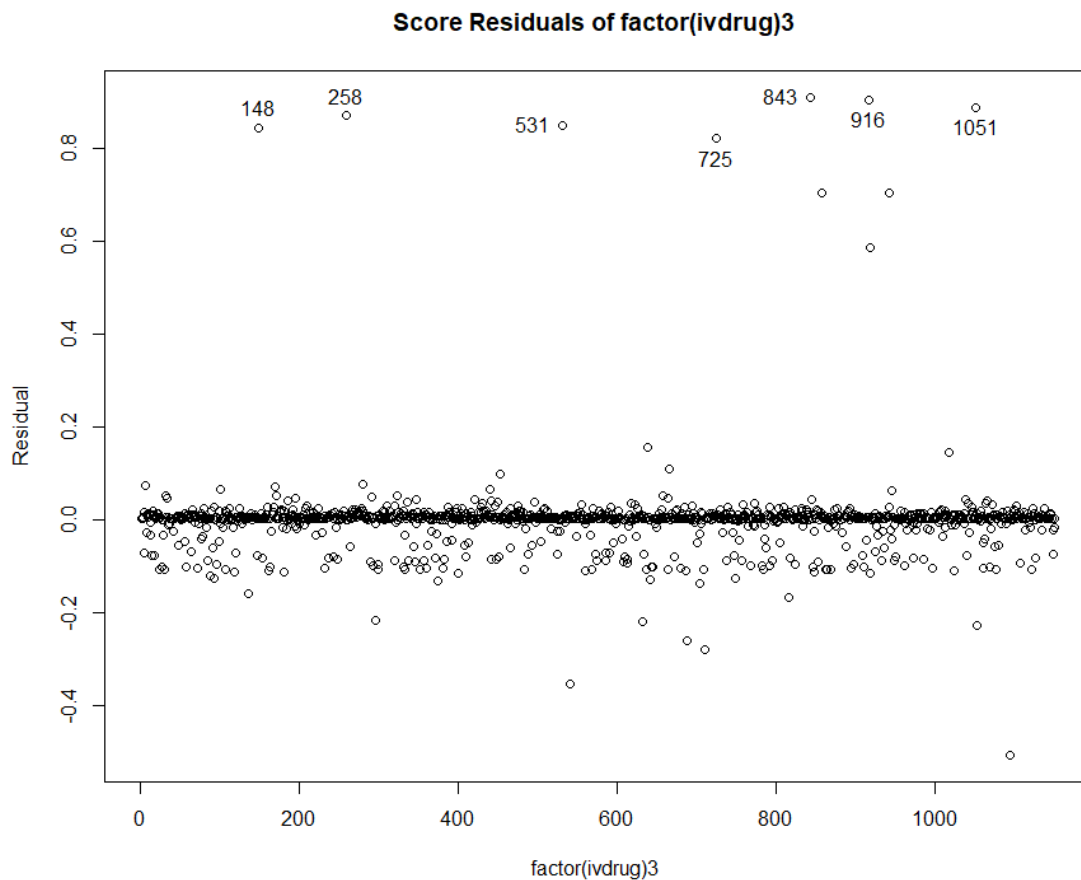

Score Residuals for Karnofsky Scores

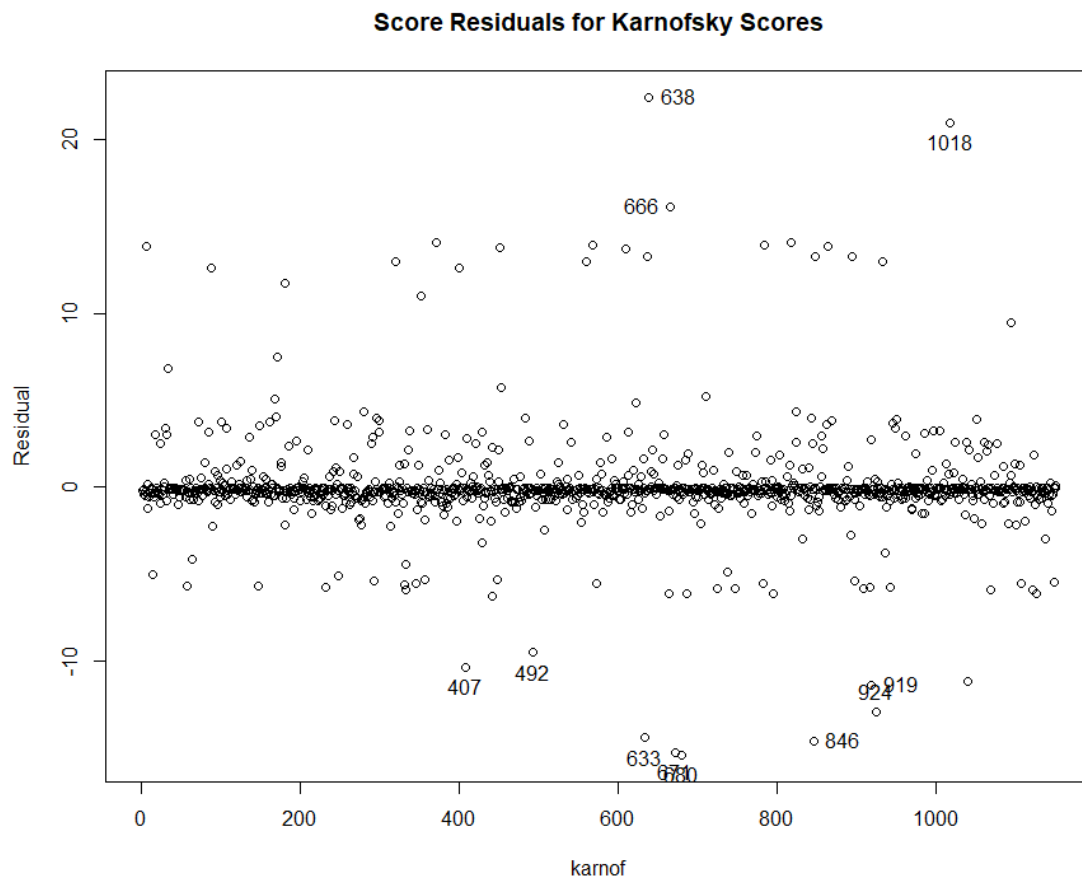# Figure 11: Score Residuals for $\sqrt{cd4}$

Score Residuals for sqrt(cd4)
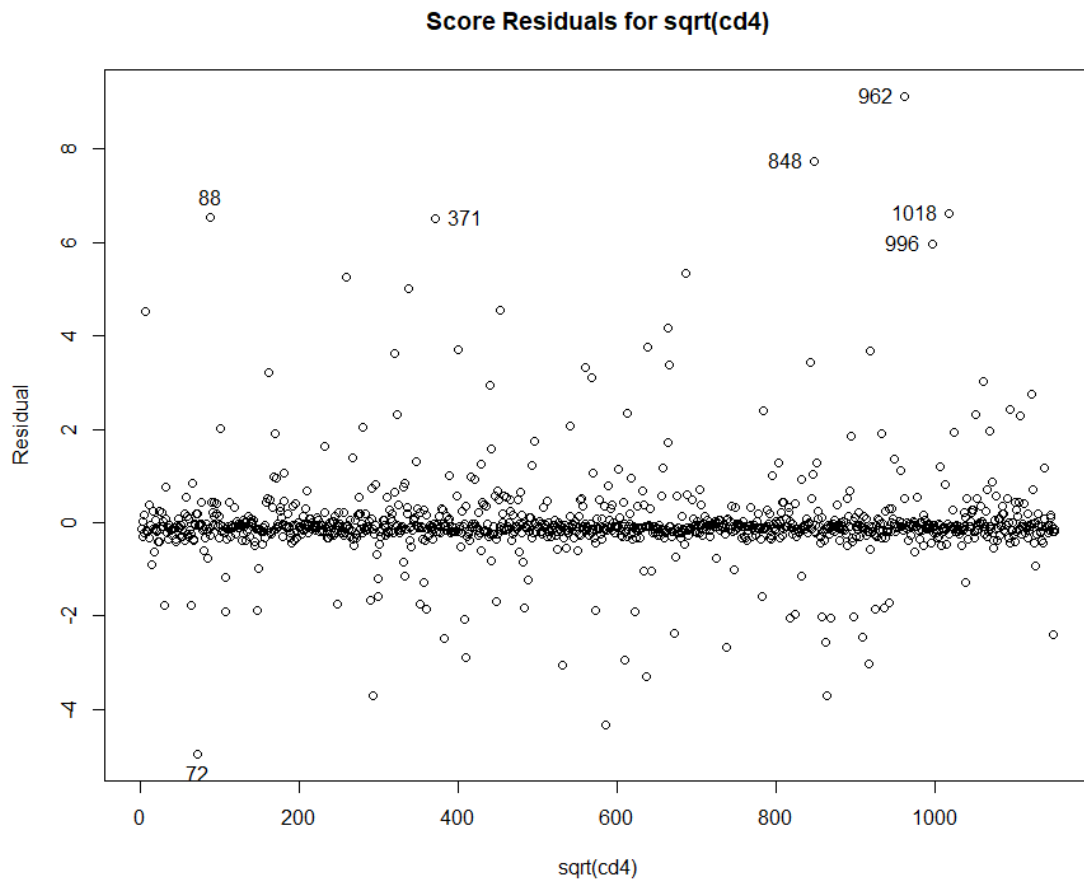
Score Residuals for I(age^2)
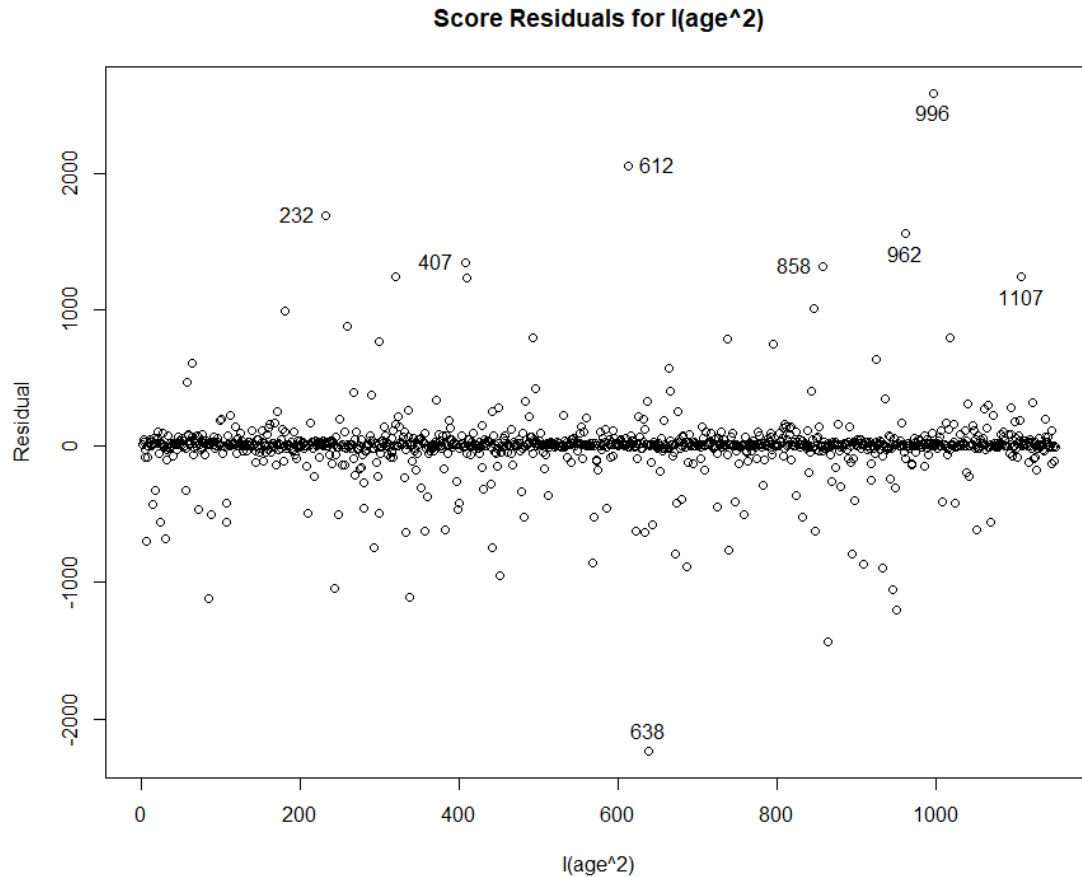
Outliers

Table 1: Outliers

	id	time	censor	tx	txgrp	strat2	ivdrug	karnof	cd4	age
492	495	144	1	1	2	1	2	70	57	56
1038	1,043	221	0	0	1	1	2	80	71.500	47
638	641	290	0	0	1	0	1	70	6	58

Conclusion

Figure 13: Cox PH Model Stratified by Treatment
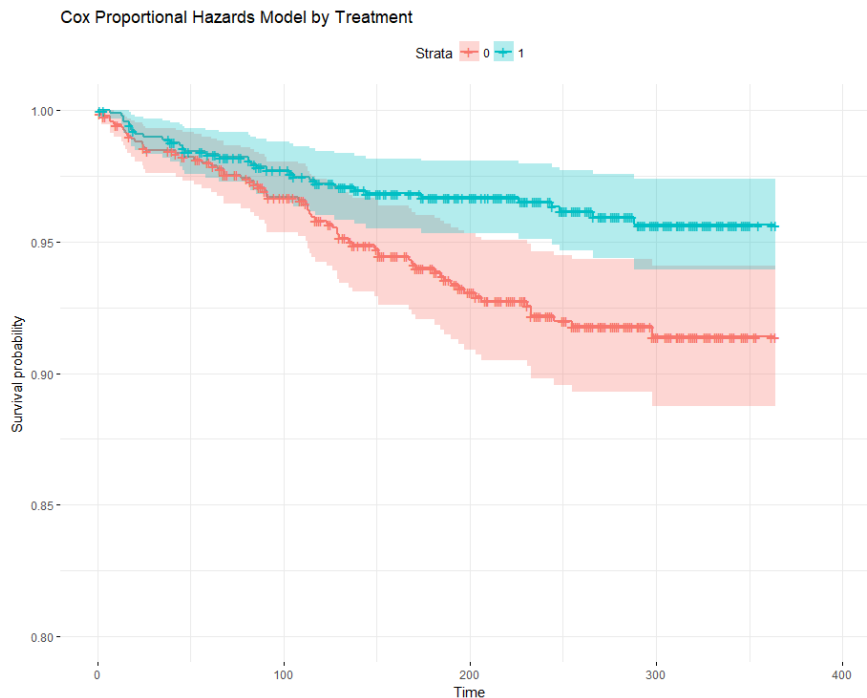
Treatment including IDV improve the survival rates than treatment without IDV.

Conclusion

IDV is no longer recommended for use in United States because of serious side effects like kidney failure and liver problems [4].

Table 2: Final Cox PH Model with Stratified Treatment with age^2

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
factor(ivdrug)2	0.818	2.267	1.022	0.801	0.423
factor(ivdrug)3	-0.600	0.549	0.336	-1.786	0.074
karnof	-0.054	0.947	0.012	-4.512	0.00001
sqrt(cd4)	-0.182	0.833	0.030	-6.132	0
I(age^2)	0.0002	1.000	0.0001	1.948	0.051

Alternate model with *age*

Table 3: Final Cox PH Model with Stratified Treatment and age

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
factor(ivdrug)2	0.820	2.270	1.022	0.802	0.422
factor(ivdrug)3	-0.609	0.544	0.336	-1.815	0.070
karnof	-0.054	0.947	0.012	-4.515	0.00001
sqrt(cd4)	-0.182	0.834	0.030	-6.119	0
age	0.021	1.021	0.011	1.856	0.063

Alternate model with *age* and age^2

Table 4

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
factor(ivdrug)2	0.822	2.276	1.022	0.805	0.421
factor(ivdrug)3	-0.594	0.552	0.337	-1.760	0.078
karnof	-0.054	0.947	0.012	-4.514	0.00001
sqrt(cd4)	-0.182	0.833	0.030	-6.136	0
age	-0.012	0.988	0.069	-0.179	0.858
I(age^2)	0.0004	1.000	0.001	0.487	0.627

```
Analysis of Deviance Table
 Cox model: response is  Surv(time, censor)
 Model 1: ~ strata(tx) + factor(ivdrug) +
 karnof + sqrt(cd4) + age + I(age^2)
 Model 2: ~ strata(tx) + factor(ivdrug) +
 karnof + sqrt(cd4) + age
    loglik   Chisq Df P(>|Chi|)
1 -547.70
2 -547.81  0.2274   1      0.6334
```

```
Analysis of Deviance Table
 Cox model: response is  Surv(time, censor)
 Model 1: ~ strata(tx) + factor(ivdrug) +
 karnof + sqrt(cd4) + age + I(age^2)
 Model 2: ~ strata(tx) + factor(ivdrug) +
 karnof + sqrt(cd4) + I(age^2)
    loglik   Chisq Df P(>|Chi|)
1 -547.70
2 -547.71 0.0315   1     0.8591
```

Discussion

Future Directions for HIV Treatment Research:

- Long-Acting Drugs
- Side Effects
- Broadly Neutralizing Antibodies
- Therapeutic HIV Vaccines

Any questions? Thank you!

Bibliography

📄 "Hiv/aids," Mar 2018. [Online]. Available: https://www.cdc.gov/hiv/basics/whatishiv.html

📄 J. Levin, Aug 1998. [Online]. Available: http://www.natap.org/1998/drug/320.html

📄 L. Xu, "Assessing the fit of the cox model."

📄 L. J. Campbell, F. Ibrahim, M. Fisher, S. G. Holt, B. M. Hendry, and F. A. Post, "Spectrum of chronic kidney disease in hiv-infected patients." Jul 2009. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19226409