



# Survival Analysis approach in evaluating the efficacy of ARV treatment in HIV patients at the Dr GM Hospital in Tshwane, GP of S. Africa

Marcus Motshwane  
Dept. of Statistics  
University of Limpopo  
Pretoria-S.A.



# Background

- Survival analysis is aimed at estimating the probability of survival, relapse or death that occurs over time
- Relevant in clinical studies evaluating the efficacy of treatments in humans or animals
- Commonly deals with rates of mortality and morbidity



# Problem statement

- The efficacy of ARV treatment at Dr G Mukhari hospital is favourable, but not clear as to the extent they are helpful to patients
- Survival analysis, a scientific statistical tool is conducted to-
  - model ARV treatment efficacy in HIV patients
  - confirm the association between survival or not of patients after ARV treatment

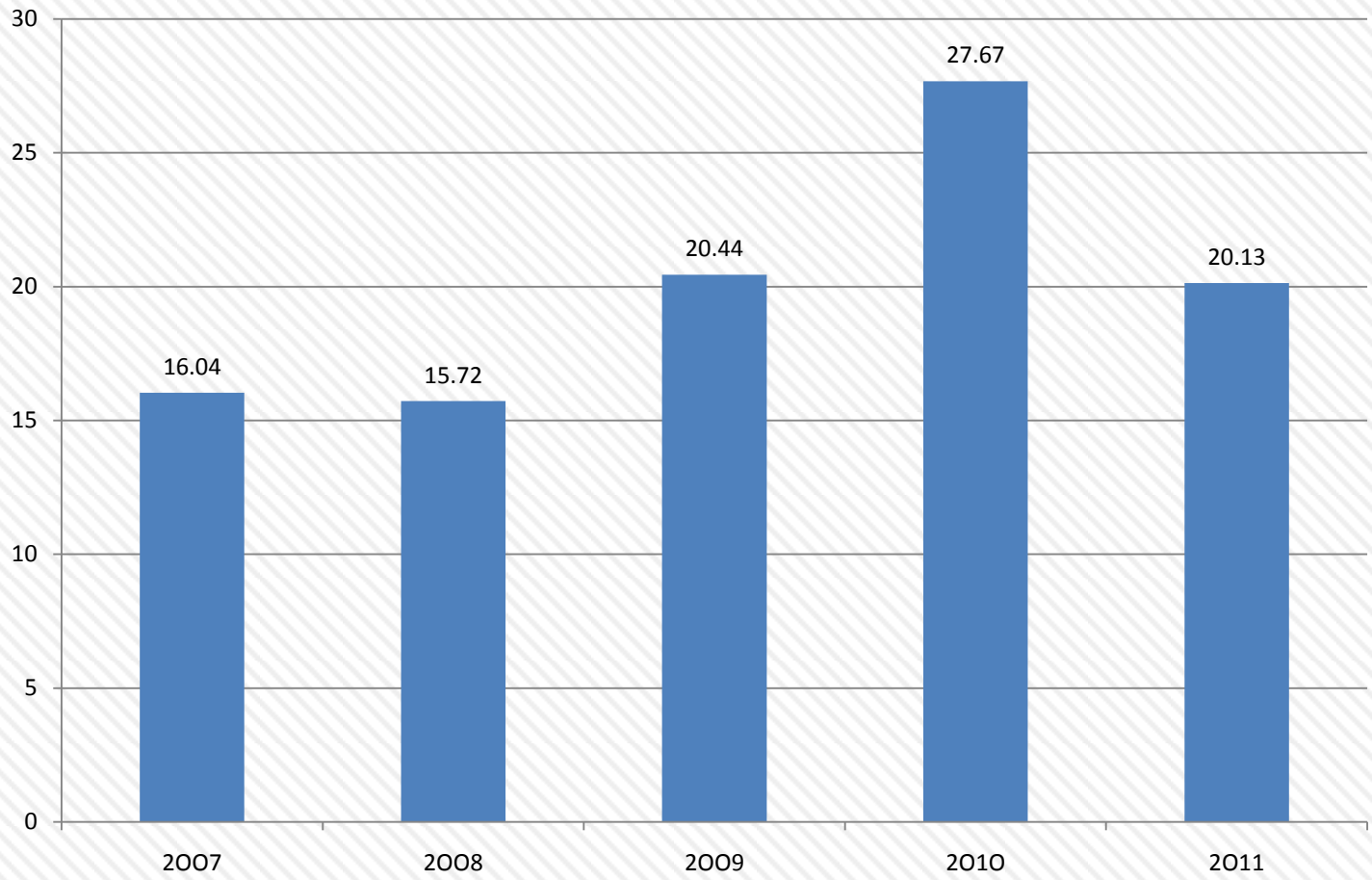


# Data Analysis

- 2007-2011 raw data
- 318 HIV/AIDS cases
- 24 variables selected
- STATA (12), SAS (9.2) & SPSS(21)

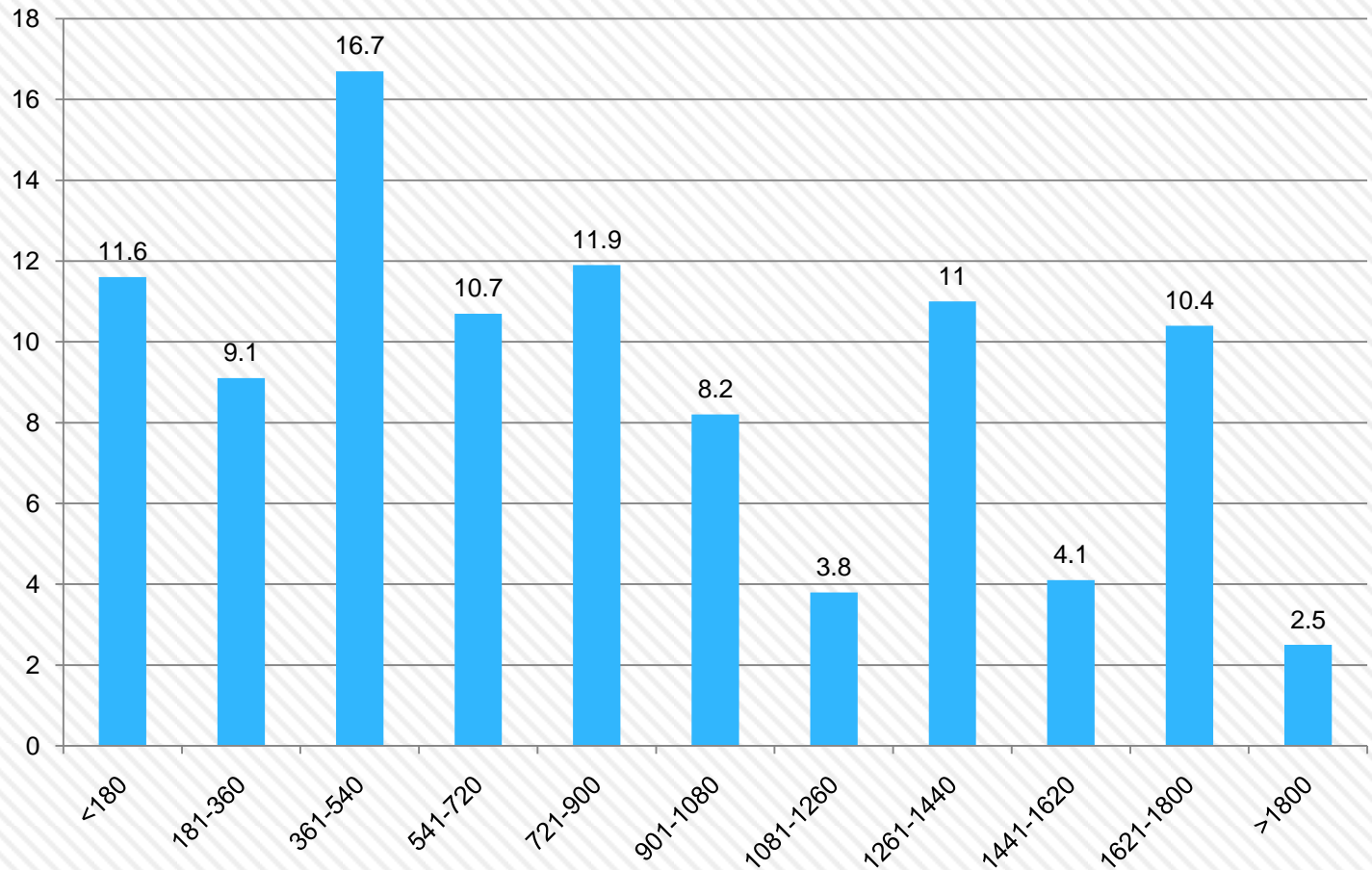


# Year



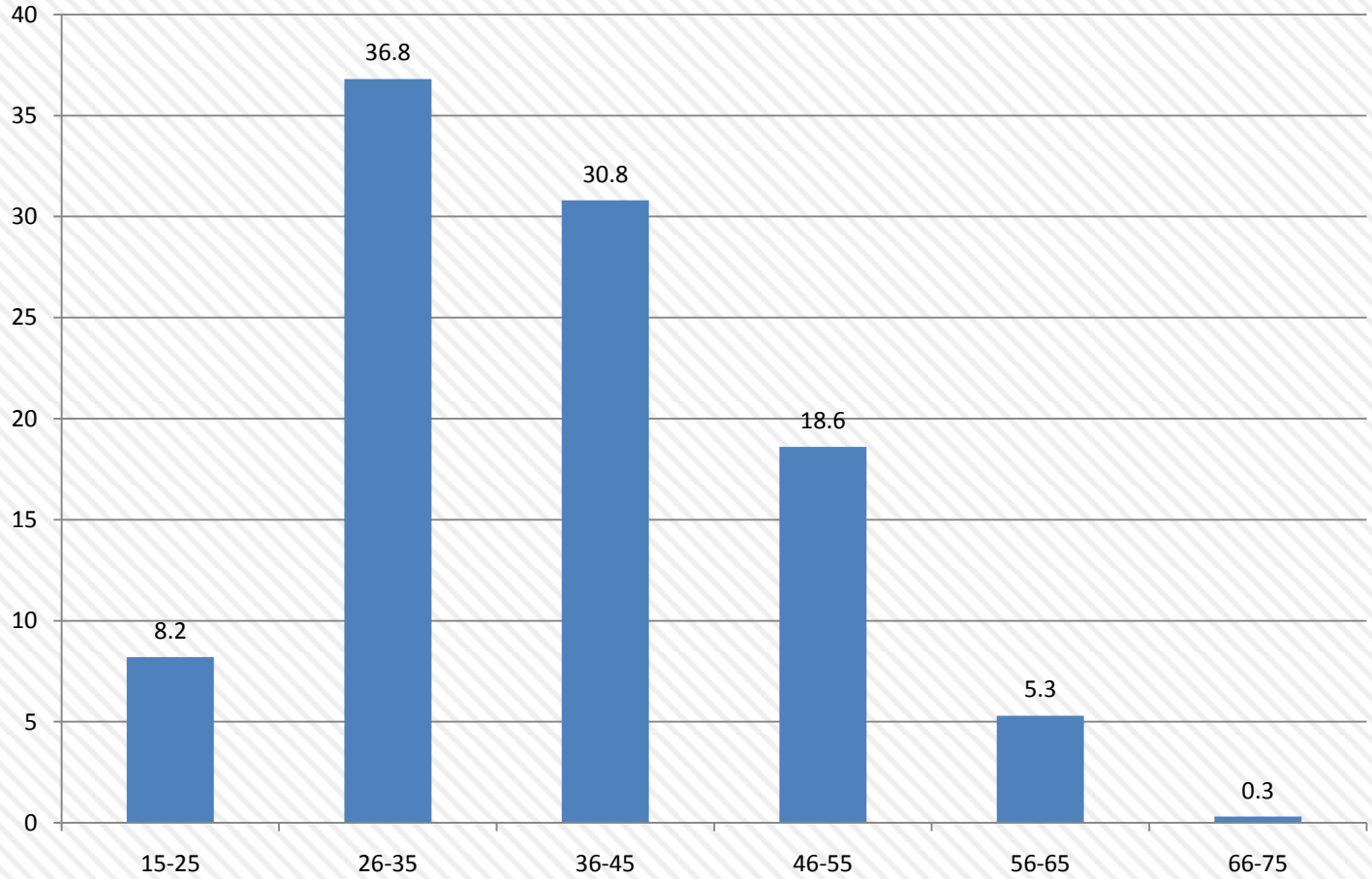


# No. of days on ARV



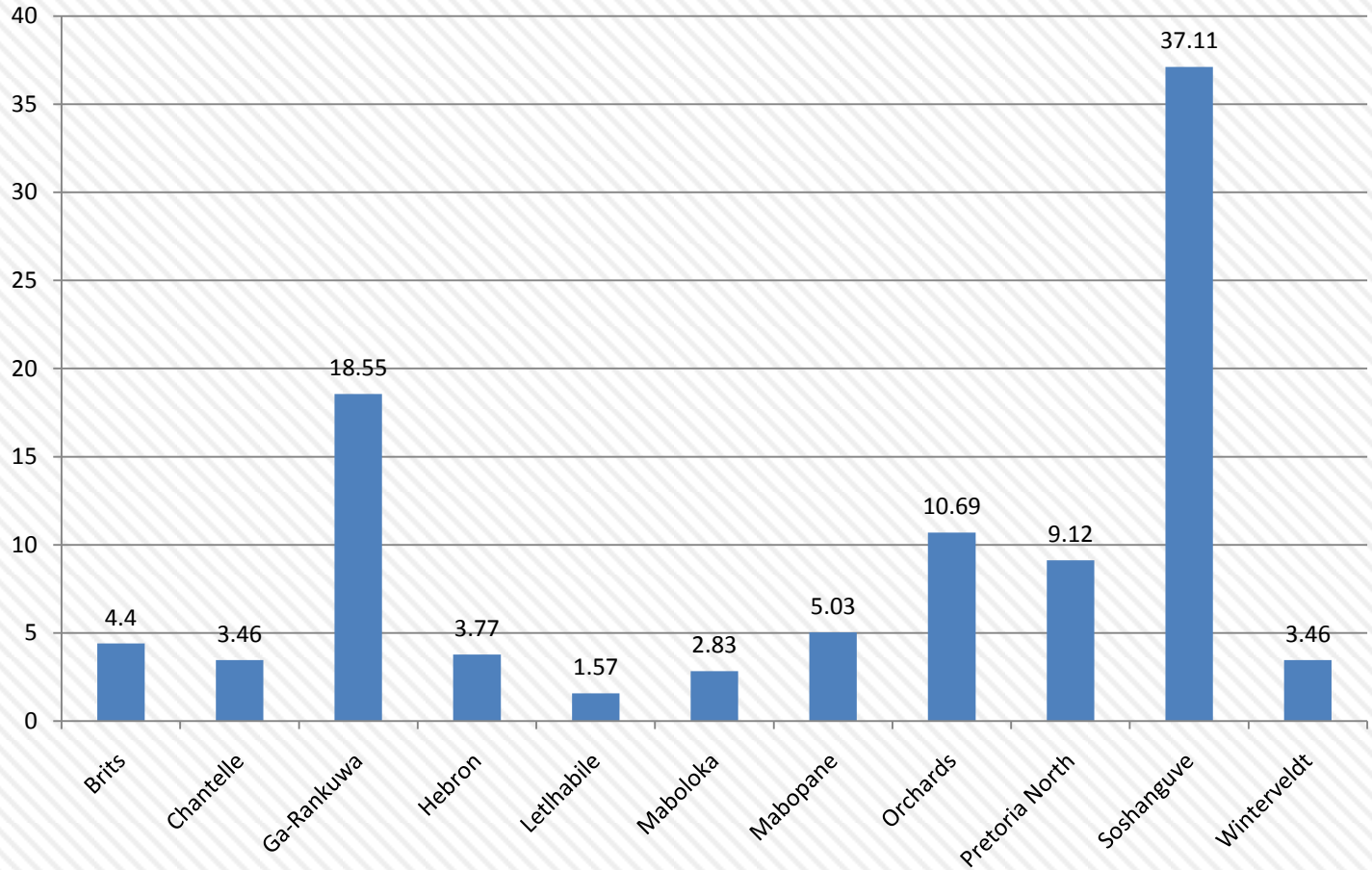


# Age (Years)





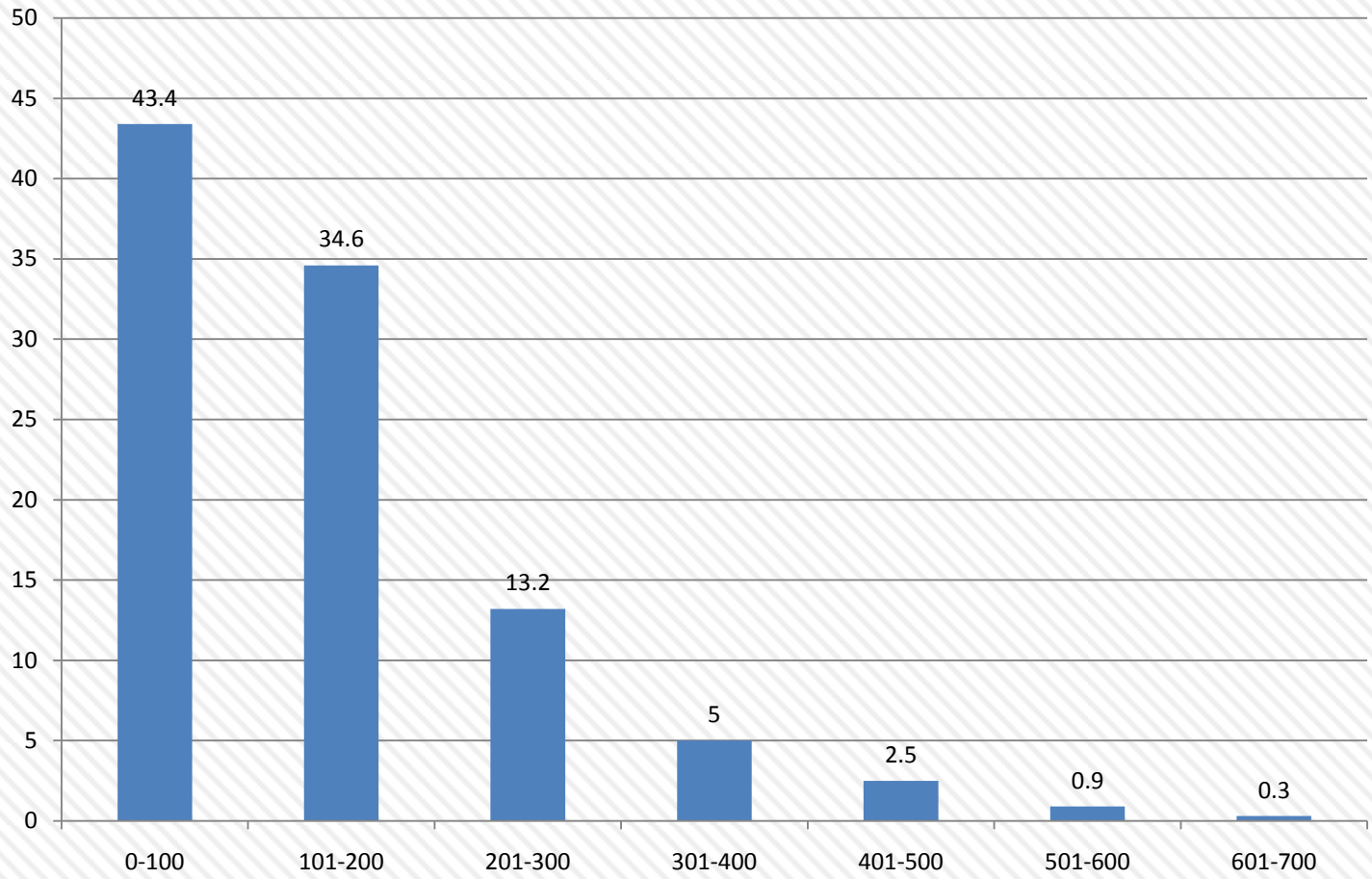
# Residential Area





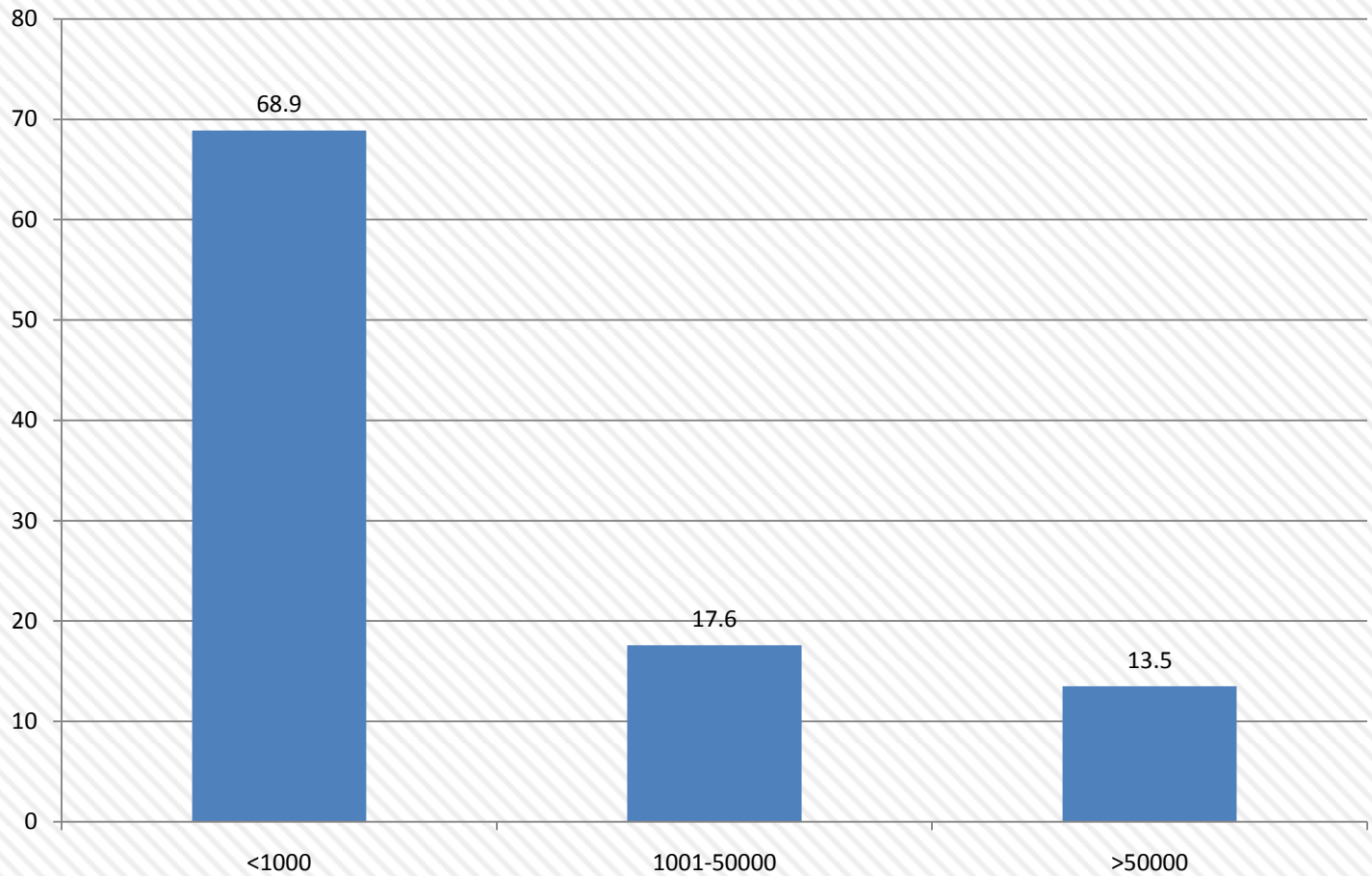


# CD4 Count





# Viral Load





# Survival function

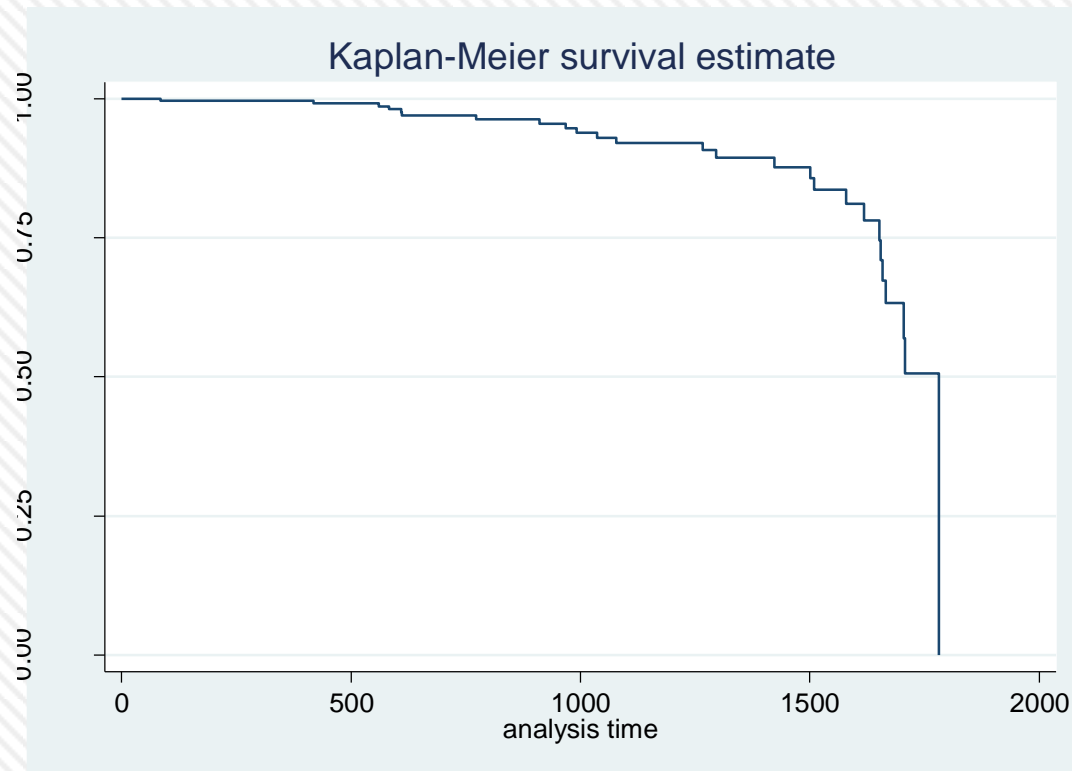
- Used to describe the time-to-event concept for all patients.
- This is the probability of an individual to survive beyond time “x” and is defined as :

$$S(x) = P(X > x)$$

- Is a non-increasing function with a value of 1 at the origin and 0 at infinity.
- This is the case here with 1 at four days (4) and zero at the end of (1781) days, meaning that there is conformity with the survival function.



# Kaplan-Meier Estimate





# Kaplan-Meier estimate (cont)

- It is an estimator of the survival function- also called the product limit estimator
- It is a function of the probability of survival plotted against time
- In the *ith* interval, the probability of death can be estimated by:

$$\frac{d_1}{n_1}$$



# Kaplan-Meier (cont)

- The estimated survival probability is:

$$\hat{S}(t) = \prod \left( \frac{n_1 - d_1}{n_1} \right)$$

*Where,*

- “d” = number of deaths observed at time “t”
- “n” = the number of patients at risk



# Kaplan-Meier (cont)

- All patients were alive at time  $t=0$
- They remained so until the first patient died after four (4) days.
- The estimate of the probability of surviving at zero is  $1.0$
- *The estimate of the survival function is thus:*  $\hat{S}(t) = 1.0$   
at  $t=0$



# The log-rank test

Log-rank test for equality of survivor functions

	Events	Events
group	observed	expected
0	0	20.09
1	26	5.91
Total	26	26.00

chi2(1) = 105.26

Pr>chi2 = 0.0000



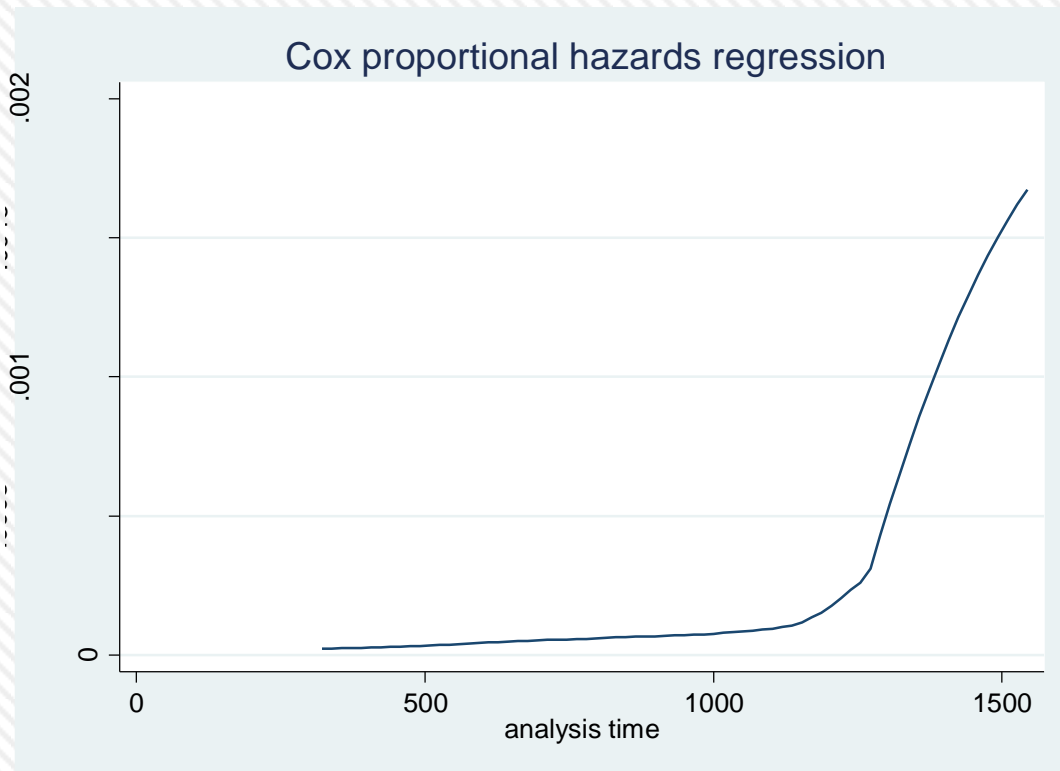


# The log-rank test (cont)

- The observed values are different from the expected values
- Produce a highly significant chi-squared value ( $P < 0.05$ ).
- The null hypothesis is rejected at the 5% level of significance
- Survivor functions of the two groups are not the same.



# Hazard function



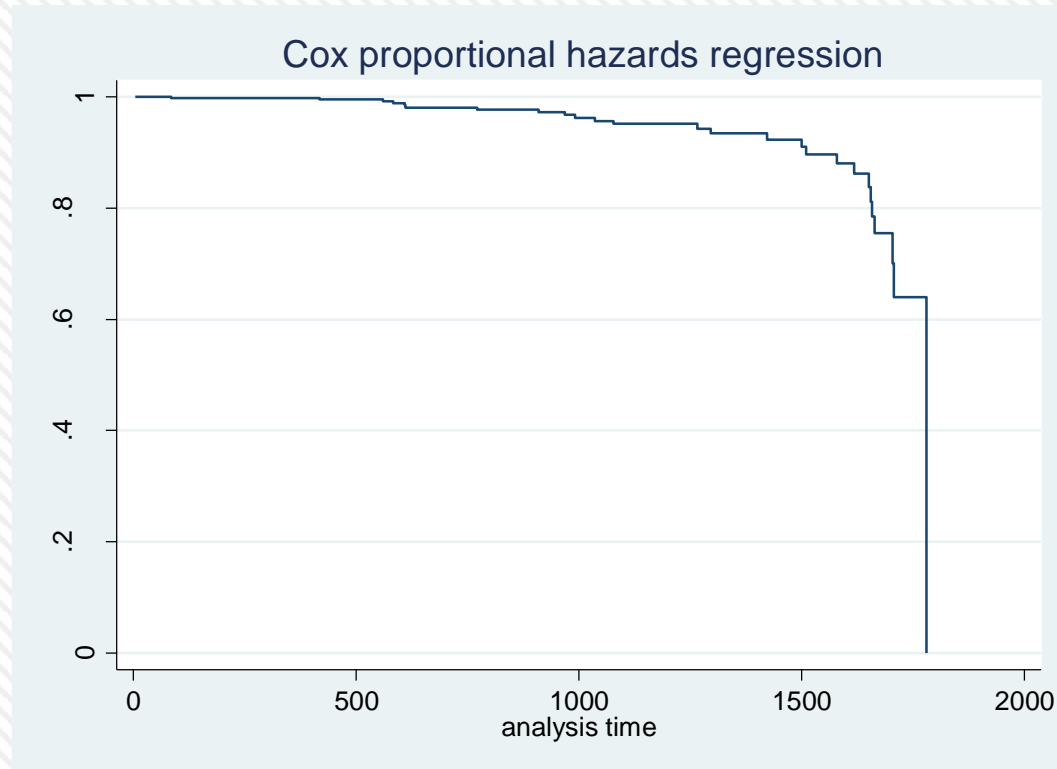


# Hazard function

- It is the proportion of subjects dying or failing in an interval per unit of time
- As days pass, the number of patients dying also increases
- It is an increasing function as opposed to the non-increasing function of the Kaplan-Meier survival estimate.



# Cox survival curve





# Cox survival curve

- The graph of the estimated baseline survivor function
- The Cox approach is the most widely used regression model in survival analysis
- The probability of survival is 1 at time  $t=0$
- Drops to “0” as the number of days elapses to maximum of 1781.



# Cox regression model

```

No. of subjects =          312                Number of obs   =          312
No. of failures =           26
Time at risk    =        249675
Log likelihood   =   -96.694567
LR chi2(7)      =          13.54
Prob > chi2     =          0.0599
  
```

_____	_____	_____	_____	_____	_____	_____
_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+	-----	-----	-----	-----	-----	-----
gender2	.6791803	.2879292	-0.91	0.361	.2958897	1.558979
age	1.032749	.0259595	1.28	0.200	.9831025	1.084903
marital2	.3615232	.1486429	-2.47	0.013	.1614947	.8093083
education2	1.084283	.0966501	0.91	0.364	.9104767	1.291268
township2	.8435348	.0977518	-1.47	0.142	.6721447	1.058628
cd4	.9987908	.0023613	-0.51	0.609	.9941735	1.00343
viral	1	5.54e-08	-0.20	0.845	.9999999	1
-----	-----	-----	-----	-----	-----	-----



# Cox reg.model (cont)

- It asserts that the hazard rate for the *ith* subject in the data is

$$h(t | x_i) = h_0(t) \exp(x_i \beta_i)$$

- The model is thus:

$$h(t | ARV) = h_0(t) \exp(0.67 \text{gender} + 1.03 \text{age} + 0.36 \text{marriage} + 1.08 \text{edu} + 0.84 \text{township} + 0.99 \text{cd4} + \text{viral})$$



# Cox reg.model (cont)

- Since  $P > 0.05$  for gender, age, education, township, c d4 count and viral load,
- No significant statistical difference amongst these variables with regard to the predictor variable, days ARV.





# Conclusion

- 92%(292/318) were alive after ARV treatment as compared to 8%(26/318) that died
- At the 5% level of significance, significant hazard ratios were characterised by hazard ratios that are significantly different from “1”, and 95% confidence interval (CI)
- ARV had a significant statistical impact on AIDS patients’ survival
- Overall mortality rates have decreased

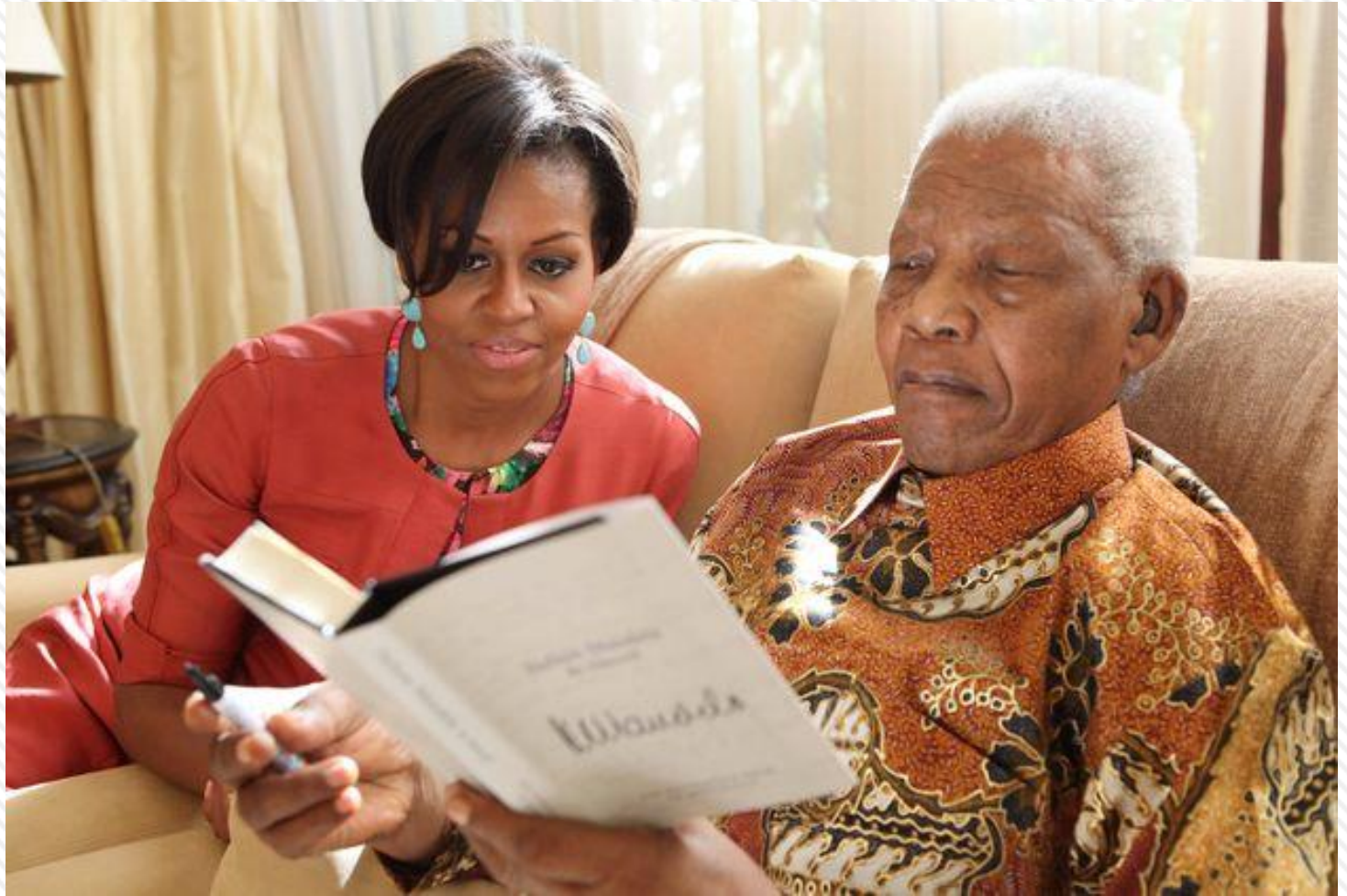


# IN MEMORY





# Mrs Obama





FINALLY

THANK YOU