CHAPTER IV

RESULTS

The Study Population

Analyses were based on the 404 individuals from the HIVNAT 006 cohort who fulfilled the inclusion criteria of never having taken any ARV medication before enrolment into their first study and who achieved virological suppression. The final cohort consisted of 221 males and 183 females, representing 553 083 patient days on study.

The first patient in this study cohort enrolled on 7 February 2000 and the last patient on 20 December 2006. The date of first virological suppression was 3 April 2000 and the date of last virological suppression was 17 September 2007. The date of last data collection was 17 December 2007. During this time, no participants were lost to follow-up.

Baseline Characteristics

Table 6 and Table 7 show the baseline (pre HAART initiation) demographic and clinical characteristics of the 404 ARV-naïve individuals who enrolled between February 2000 and December 2006 and who achieved initial virological suppression.

The Shapiro-Wilk test was used to test the null hypothesis that the samples in these variables came from a normally distributed population. The *P* values for these tests were consistently <0.05, indicating that the null hypothesis of normality should be rejected: that is, these data were not normally distributed. The results of Shapiro-

Wilk tests were for *gender*, W=0.633296, *P*<0.0001, *age* at baseline, W=0.970229, *P*<0.0001, *weight* at baseline, W=0.975639, *P*<0.0001, *CD4+ count* at baseline, W=0.921317, *P*<0.0001, and *pVL* at baseline, W=0.991931, *P*=0.0273. In the case of non-normal data, it is more appropriate to report the median with its interquartile range (IQR), that is the range from the 25th percentile to the 75th percentile, as the measure of central tendency. (Lang & Secic, 2006)

There were 221 (54.7%) males and 183 (45.3%) females. The median age at baseline was 32.7 years (IQR 28.0 to 37.9 years). The median baseline weight was 56.2 kg (IQR 50.0 kg to 64.3 kg). The mean weight for males was 62.39 kg (Standard Deviation [SD] 9.97) and for females was 52.05 kg (SD 8.56). There was a highly statistically significant difference in the weights of males and females (t=11.22, *P*<0.0001). HIV transmission was mainly through heterosexual contact, 290 (71.8%), with 89 (22.0%) individuals being infected through homosexual contact and 25 (6.2%) individuals by other routes, including injecting drug use, occupational exposure and unknown.

Most participants did not have advanced HIV disease. The Centres for Disease Control and Prevention (CDC) Classification System for HIV Infection in Adults and Adolescents were Category A 180 (44.6%), Category B 158 (39.1%), Category C 65 (16.1%), with 1 categorization missing. Briefly, CDC A is asymptomatic HIV infection, CDC B is symptomatic HIV infection and CDC C is defined as AIDS. (CDC, 1993).

This study examined CD4 T-cell lymphocyte (CD4+) count and plasma viral load (pVL) measurements at baseline and at study exit, whether at the time of virological failure or at censoring. The date of censoring was the date of the last

available laboratory assessment. Median baseline CD4+ count was 167 cells/mm^3 (IQR 57 to 261 cells/mm^3) and baseline pVL was 4.7 log_{10} copies/mL (IQR 4.3 to 5.4 log_{10} copies/mL). The majority of participants were hepatitis B and hepatitis C negative (78% and 92.6%) respectively.

Table 6: Subject characteristics

Category	Parameter	n	(%)
Gender	Male	221	54.7
	Female	183	45.3
Baseline disease status*	CDC A	180	44.6
	CDC B	158	39.1
	CDC C	65	16.1
	Missing	1	0.2
Mode of HIV transmission	Heterosexual	290	71.8
	Homosexual	89	22.0
	Other**	25	6.2
Hepatitis B status	Positive	85	21.0
	Negative	315	78.0
	Missing	4	1.0
Hepatitis C status	Positive	21	5.2
	Negative	374	92.6
	Missing	9	2.2
First regimen	NNRTI[1]-based	291	72.0
	PI[2]-based	113	28.0
Virological failure ***	Failed	69	17.1
	Not failed	335	82.9
Calendar year of entry	2000	207	51.2
	2001	16	4.0
	2002	34	8.4
	2003	13	3.2
	2004	46	11.4
	2005	76	18.8
	2006	12	3.0

* Centres for Disease Control and Prevention Classification System for HIV Infection in Adults and Adolescents, 1987/1993

** Other modes of transmission were by injecting drug use, occupational exposure, blood transfusion or missing data

*** Virological failure is defined as pVL >50 copies/mL measured on two consecutive occasions at least 4 weeks apart

1. Non-nucleoside reverse transcriptase inhibitor

2. Protease inhibitor

At baseline, 291 (72%) of this cohort commenced a NNRTI-based ARV regimen and the remaining 113 (28%) commenced a PI-based regimen. The majority of patients, 207 (51.2%), enrolled in 2000, followed by 76 (18.8%) in 2005 and 46 (11.4%) 2004. The median time on study was 1038 days (IQR 605.5 - 2256.5 days).

Table 7: Medians and interquartile ranges

Category	Measure	(n %)	Median	IQR†
Baseline age	Years	404 (100)	32.7	28.0 to 37.9
Baseline weight	Kilograms	404 (100)	56.2	50.0 to 64.3
Baseline CD4+ count	cells/mm^3	404 (100)	167	57 to 261
Baseline pVL	log_{10}RNA	404 (100)	4.7	4.3 to 5.4
CD4+ count at exit	cells/mm^3	403 (99.8)	462	314 to 630
pVL at exit	log_{10}RNA	404 (100)	1.7	1.7 to 1.7
Time on Study	Days	404 (100)	1038	605.5 to 2256.5
Time to virological suppression	Days	404 (100)	119	85.0 to 189.5
Time to failure following suppression	Days	69 (17.1)	954	842 to 1085

† Interquartile range

Time to Virological Suppression

All 404 subjects in this analysis achieved initial virological suppression. Median time to virological suppression after commencement of HAART was 119 days (IQR 85–190 days).

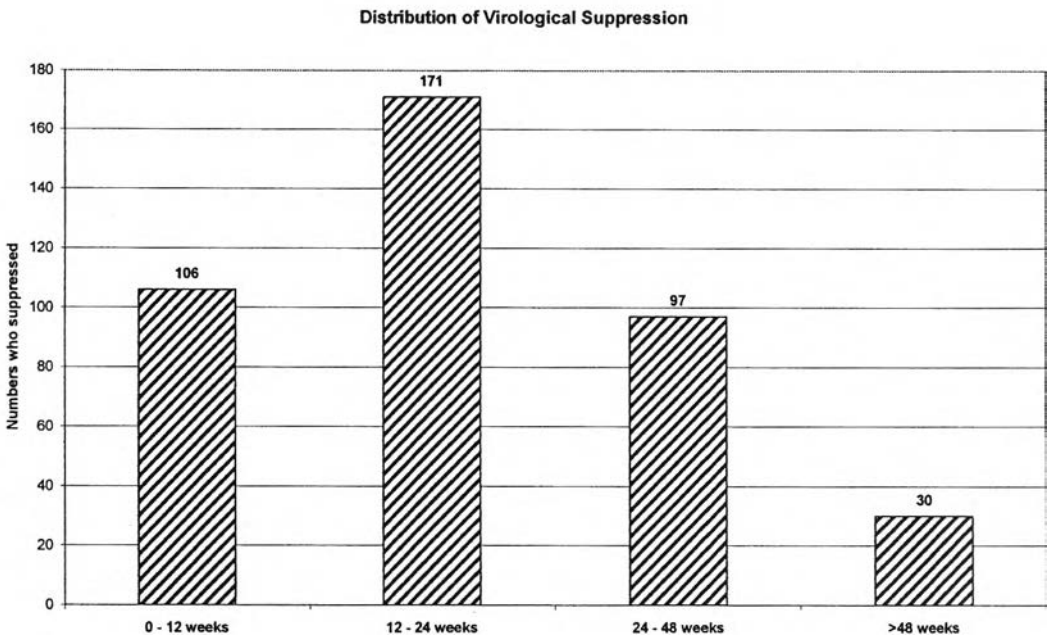
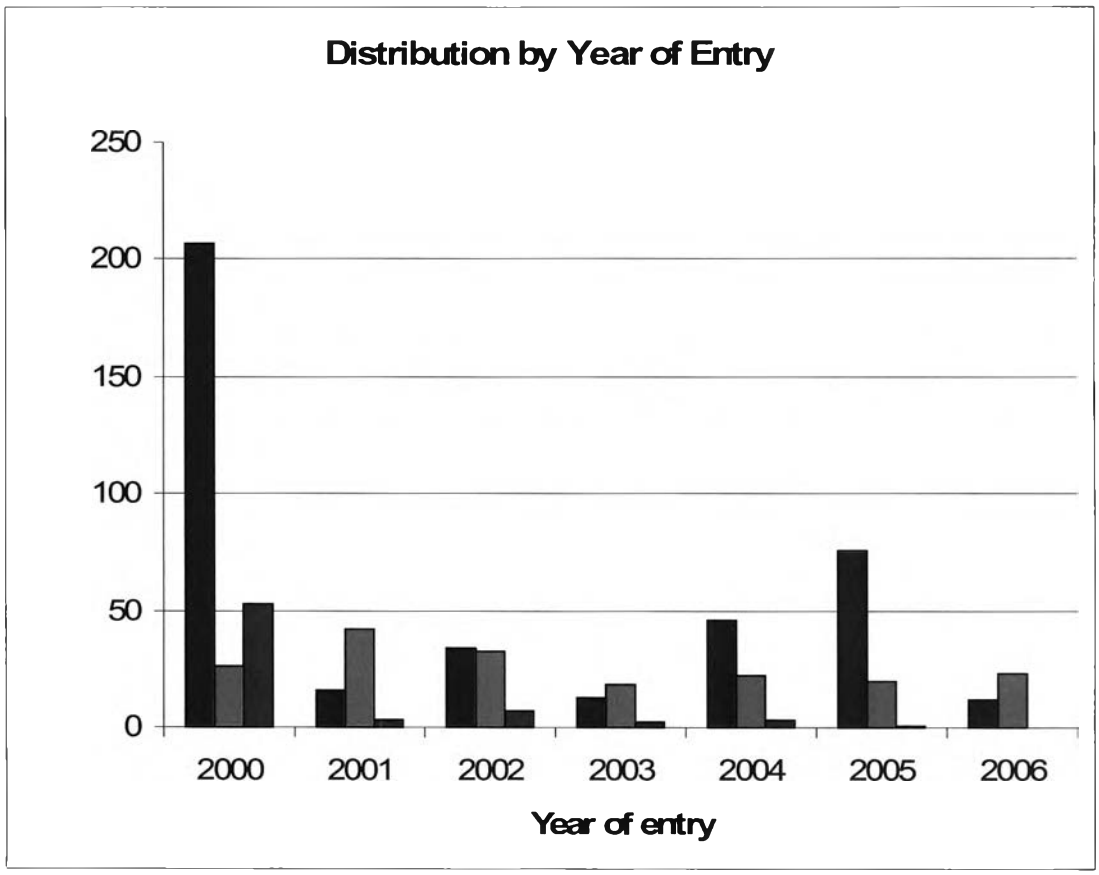


Figure 3: Time to virological suppression

Figure 3 shows the time to virological suppression in weeks. Most subjects achieved virological suppression early, 171 (42.3%) in weeks 12 to 24 and 106 (26.2%) in weeks 0 to 12. Although the analyses were conducted using person days, the data are displayed in weeks as this is the most widely used type of presentation. Also, at HIV-NAT, clinic visits typically are scheduled in periods of twelve weeks. The Kaplan-Meier curve shows the same effect (Figure 3).



Number enrolled	207	16	34	13	46	76	12
Mean wks to supp.	26	42	32	18	22	20	23
Failures	53	3	7	2	3	1	0

Figure 4: Distribution by year of entry onto study

Figure 4 shows the distribution by calendar year of entry into study, virological failures and mean weeks to suppression by year of entry. It should not be inferred from this figure that a longer time on study was associated with fewer failures. As can be seen, most of the patients (51.24%) enrolled in 2000 and 76.8% failed in the same year. These are unadjusted data, presented for descriptive purposes only.

Independent variables (covariates)

Based on previous studies, eleven candidate explanative characteristics, modelled with independent variables, were assessed: baseline measurements for gender, age, weight, CD4+ count, pVL, ARV regimen, co-infection with hepatitis B and/or C, clinical stage of HIV disease (CDC A, B or C), mode of transmission and year of entry. (Dragsted et al., 2004; Fournier et al., 2005; Manegold et al., 2004)

Initially, the variable for time was modelled. Bivariate assessment of characteristics was conducted and variables were selected for multivariable modelling. The model was fit by using person-time logistic regression and its consistency was evaluated by supplemental Poisson regression.

Results for Time from Start of HAART to Virological Suppression

Person-time logistic regression was used to determine odds ratios for various independent variables that might have been associated with time from the start of HAART to virological suppression in this study. For the interval time to virological suppression, the original data set of the 404 eligible patients was expanded with a SAS "do-loop" to 71 113 person days. The median time to suppression was 119 (IQR 85.0 to 189.5) days.

The Kaplan-Meier survival distribution function was used in the unexpanded data set to visually assess time to virological suppression (Figure 5). This curve starts at zero; nobody suppressed at time zero. The curve shows two inflection points, suggesting that modelling time as a cubic polynomial would achieve a better logistic model than would modelling time as a linear (first-order) function or quadratic polynomial. The curve displays a relatively slow initial rise, followed by rapid

achievement of virological suppression up until about day 250, and then a much slower rate of suppression until all are suppressed.

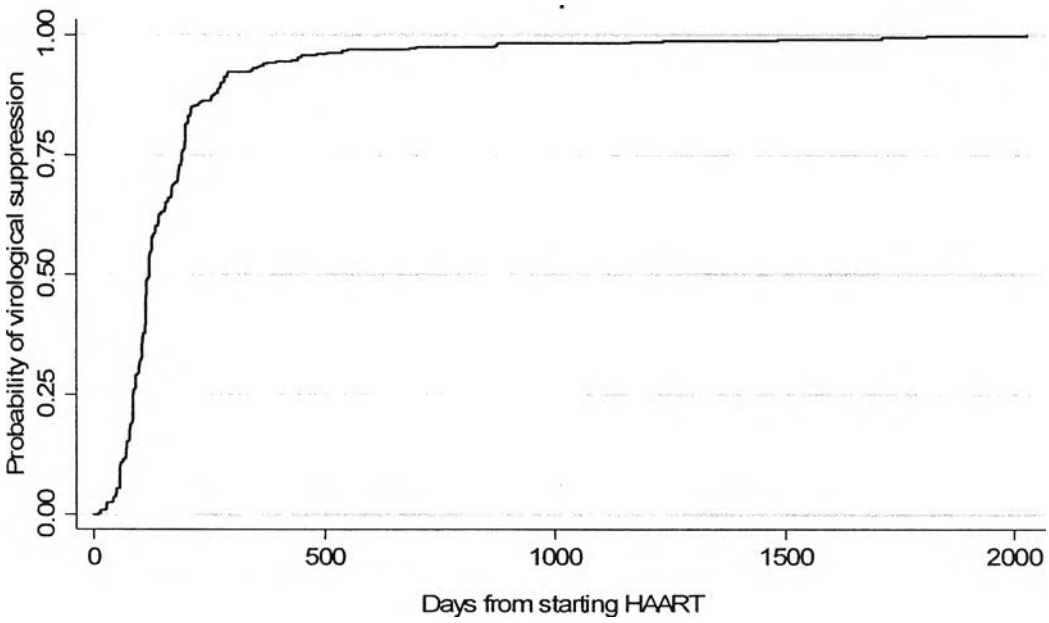


Figure 5: Kaplan-Meier curve: probability of virological suppression after starting HAART

Modelling Time

A polynomial transformation is useful as it can accommodate a wide variety of different relationships between x and y in the linear equation, and can result in an improvement over an untransformed linear function (Allison, 1999) Polynomials are constructed by putting integer powers of x in the regression model. Whenever a higher power is added to the equation, all the lower order powers must be included in the model. Although further polynomial transformation is possible, it is not customary. (Allison, 1999)

The quadratic equation is: $y = A + B_1x + B_2x^2$

The cubic equation is: $y = A + B_1x + B_2x^2 + B_3x^3$

The variable for time achieved by the expansion of the data set was modelled alone using logistic regression. The Likelihood Ratio χ^2 of the model was 11.62, P=0.0007 with only the first-order time on study. When modelled by itself including quadratic transformation of the time variable, the likelihood ratio χ^2 of the model was 25.40, P<0.0001, a clear indication of a better fit of the model. The addition of cubic transformation of the time variable produced an even better fit, the likelihood ratio χ^2 of the model was 124.5334, P<0.0001.

Bivariate Logistic Regression: from Start of HAART to Virological Suppression

The time variable with quadratic and cubic transformation was then modelled in bivariate analysis against 11 covariates: gender, baseline weight, mode of transmission (always modelled with two dummy variables, homosexual vs. heterosexual and other vs. heterosexual), baseline CDC category (AIDS vs. categories 1 and 2), baseline age, year of entry, hepatitis B and C infection, CD4+ count, pVL and baseline regimen.

Table 8: Results of bivariate logistic regression from start of HAART to virological suppression

Independent variable time	Likelihood ratio	*P* value of the model
Model with time and no other independent variables†	**124.5334**	**<0.0001**

† The time variable includes quadratic and cubic transformation

Independent variables		OR	(95% CL)	*P* value
Gender	female vs. male	1.270	(1.041 to 1.548)	0.0183*
Weight at baseline	per kg	0.998	(0.989 to 1.007)	0.6969
Mode of transmission	homosexual vs. heterosexual	0.911	(0.717 to 1.157)	0.4455
	other vs. heterosexual	1.268	(0.840 to 1.915)	0.2591
Baseline CDC category		0.798	(0.697 to 0.913)	0.0010*
Age at entry	in years	0.993	(0.979 to 1.006)	0.2910
Year of entry	from 2000 to 2006	1.041	(0.994 to 1.089)	0.0878*
Hepatitis B status	positive vs. negative	0.916	(0.720 to 1.165)	0.4747
Hepatitis C status	positive vs. negative	0.802	(0.495 to 1.298)	0.3687
Baseline CD4+ (cells/mm^3)	CD4+ ≥200 vs. <200	1.181	(0.968 to 1.440)	0.1005*
PVL at baseline	>50,000 copies/mL	0.716	(0.587 to 0.872)	0.0009*
Baseline regimen	.0=NNRTI, 1=PI	1.825	(1.458 to 2.284)	<0.0001*

** selected for multivariable analysis*

OR = Odds ratio, CL = Wald Confidence Limits

Table 8 presents the results of bivariate assessment of covariate effects using logistic regression, from the start of HAART to virological suppression (including first-order, quadratic and cubic terms for time [days] on study, as previously described). Each independent variable characteristic was assessed separately.

In these analyses, six characteristics exhibited a P value ≤0.2. These were included in a multivariable logistic model: *gender* [Odds Ratio (OR) 1.270, 95% Wald Confidence Limits (95% CL) 1.041-1.548, *P*=0.0183]; *baseline CDC severity*

category [OR 0.798, 95% CL 0.697-0.913, *P*=0.001]; *year of enrolling into the first study* [OR 1.041, 95% CL 0.994-1.089, *P*=0.0878]; *CD4+ count at baseline* [OR 1.181, 95% CL 0.968-1.440, *P*=0.1005]; *pVL at baseline* [OR 0.716, 95% CL 0.983-2.573, *P*=0.0586]; and *ARV regimen at baseline* [OR 1.825, 95% CL 1.458-2.284, *P*<0.0001].

Multiple Logistic Regression: Start of HAART to Virological Suppression

The first step in multiple logistic regression was to model time with its polynomial transformations (as described previously) alone and then with the other selected covariates, to test the differences in the β estimates (Table 9). All *P* values for the Wald χ^2 were <0.0001. The variables remain robust, the Wald χ^2 remains highly significant, the β estimates do not change direction and their percent changes were small.

Table 9: Results of β estimates for time alone and with covariates in multiple logistic regression

Variable	β before	β after	% change
Days on study	0.00959	0.0109	12.0%
Days on study squared	-0.00002	-0.00002	0.0%
Days on study cubed	8.124E-9	8.775E-9	7.4%

The six variables identified from the bivariate analyses then were modelled with the time variable and its quadratic and cubic polynomials using multiple logistic regressions (Table 10).

In the multivariable model for time to virological suppression, the χ^2 Likelihood Ratio of the model increased from 124.5334 to 162.0201 (*P*<0.0001),

which means that the overall model χ^2 highly significantly increased by 37.4867 for the addition of the six independent variables ($P<0.0001$).

Table 10: Results of multiple logistic regressions from start of HAART to virological suppression

		OR	(95% CL)*	*P* value
Days on study		1.011	(1.009 to 1.013)	<0.0001
Days on study squared		1.000	(1.000 to 1.000)	<0.0001
Days on study cubed		1.000	(1.000 to 1.000)	<0.0001
Gender	female vs. male	1.147	(0.928 to 1.418)	0.2053
Baseline CDC	CDC C (AIDS) vs. others	0.865	(0.732 to 1.020)	0.0854
Year of entry	from 2000 to 2006	1.004	(0.952 to 1.058)	0.8945
Baseline CD4+ (cells/mm^3)	CD4+ ≥200 vs. <200	0.888	(0.690 to 1.143)	0.3556
Baseline pVL	$\log_{10}$ copies/mL	0.713	(0.564 to 0.901)	0.0046
Baseline regimen	0=NNRTI, 1=PI	1.872	(1.454 to 2.408)	<0.0001

Likelihood Ratio χ^2 of the model with 9 degrees of freedom is 162.0201 ($p<0.0001$)

* OR = Odds ratio; CL = Wald Confidence Limits

Multivariable analysis of the time to virological suppression identified two of the independent variables (other than the time variables) as statistically significant in predicting virological suppression. A baseline pVL >50,000 copies/mL was significantly more predictive of slower virological suppression compared with a pVL ≤50,000 copies/mL (OR 0.713, 95% CI 0.564 to 0.901, P=0.0046) (Figure).

A baseline ARV regimen that included a PI was highly significantly more predictive of faster virological suppression compared with a NNRTI-based regimen [OR 1.872, 95% CI 1.454 to 2.408, $P<0.0001$] (Figure 7).

In addition, baseline CDC Category C (AIDS) at baseline showed a trend to predicting slower virological suppression compared with CDC Category A and

Category B; this was marginally significant (OR 0.865, 95% CI 0.732 to 1.020, *P*=0.085) (Figure 8).

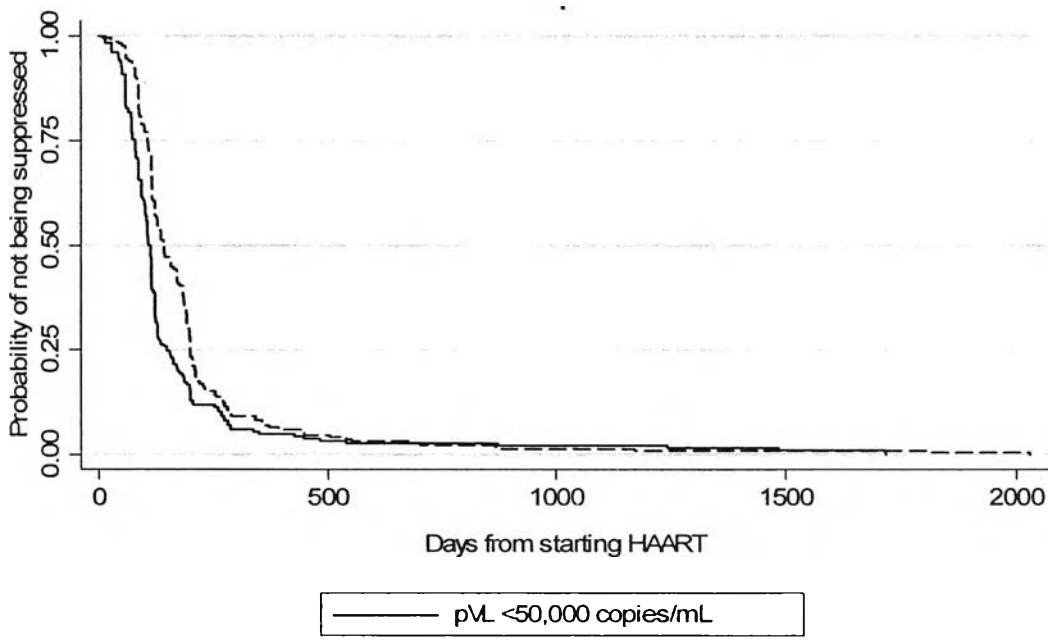


Figure6: Kaplan-Meier curve: virological suppression stratified on baseline viral load

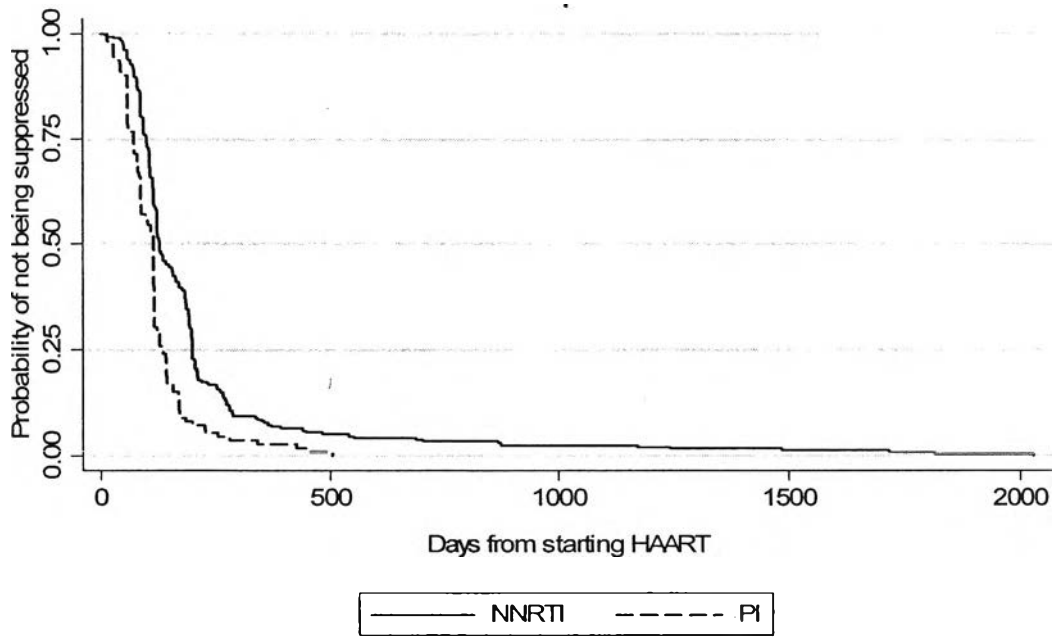


Figure 7: Kaplan-Meier curve: virological suppression stratified on baseline regimen

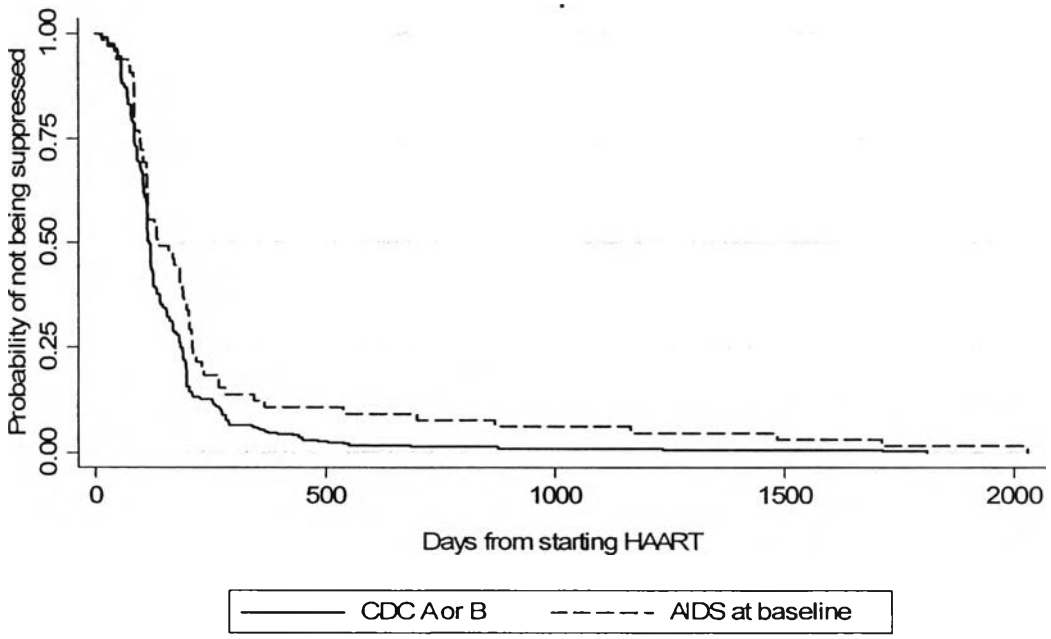


Figure 8: Kaplan-Meier curve: virological suppression stratified on CDC Category

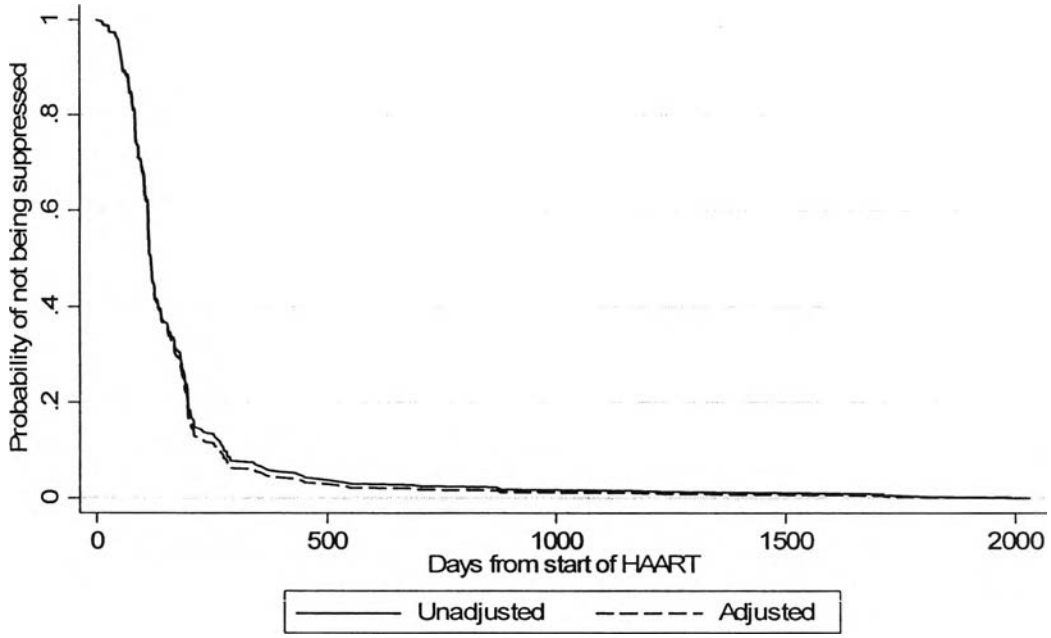
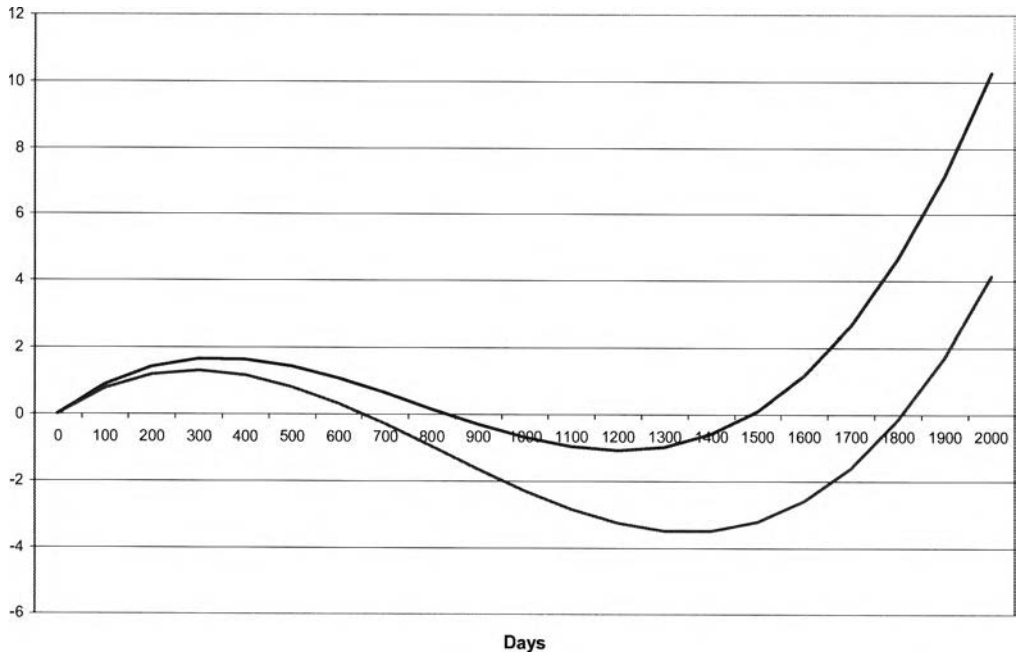


Figure9: Overlay of survival curves with the time variable unadjusted and adjusted

Figure 9 shows the overlay of the two survival curves for time in the interval to suppression. The *x*-axis shows days to suppression from the start of therapy. It is clear to see that the shapes of the curves are very similar, showing that the polynomial transformation does not distort the shape. However, as demonstrated above, the likelihood ratio χ^2 of the transformed model provides a better fit.

The log odds of these polynomial transformations were plotted as a further demonstration of the appropriateness of this technique. The equation for the curves of log odds by days of follow-up is $\beta_1(\text{days}) + \beta_2(\text{days}^2) + \beta_3(\text{days}^3)$. Figure shows the plot of the unadjusted and adjusted curves of the polynomial function of the time variable using log odds for time to virological suppression.



The green (lower) line is the unadjusted time variable, blue line the adjusted time variable

Figure 10: Time to virological suppression plotting log odds by days of follow-up

Time from Virological Suppression to Virological Failure

Sixty-nine (17.1%) of the participants in this study experienced virological failure during follow-up. The remaining 335 (82.9%) did not fail within the period of this study (Figure 11). Median time to virological failure following virological suppression was 954 days (IQR 842 to 1085 days). Median CD4+ cell count at study exit was 462 cells/mm^3 (IQR 314–630) and median exit pVL was 1.7 $\log_{10}$ copies/mL (IQR, 1.7–1.7 $\log_{10}$ copies/mL).

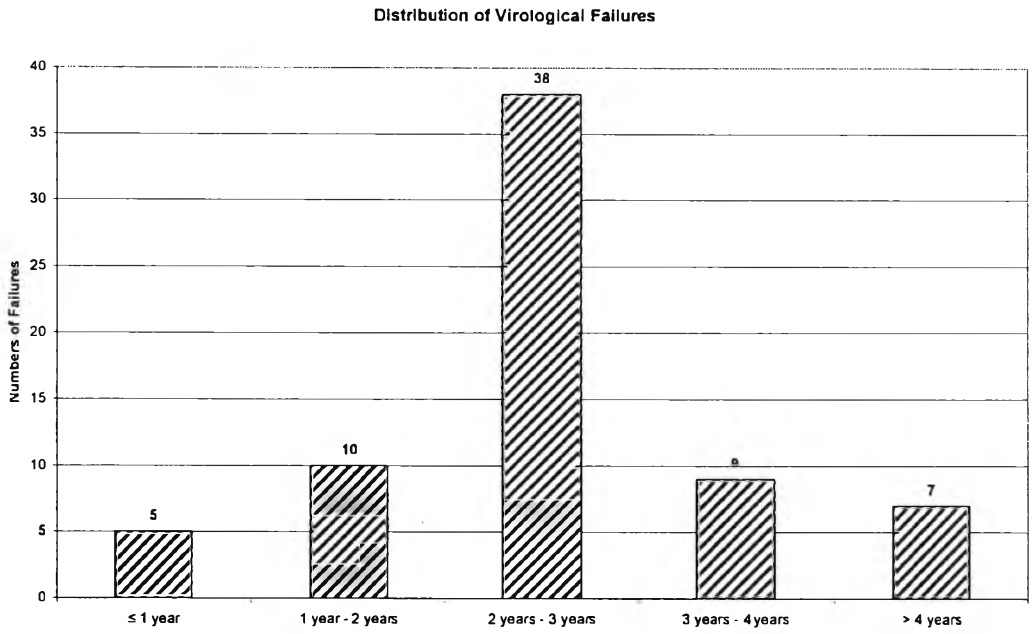


Figure 11: Distribution of the 69 virological failures

Figure 11 shows that 38 (55.1%) of the 69 subjects who failed virologically in the period of this study failed in the third year of study. Before year three, 23.2% failed and after year three, 21.7% failed, 43.9% of all failures.

Results for Time to Virological Failure

Person-time logistic regression was used to model ORs for various independent variables that might have been associated with time to virological failure.

For the interval time to virological failure, the original data set of the 404 eligible patients was expanded to 553 083 person days.

The Kaplan-Meier survival distribution function was used to describe the interval from virological suppression to virological failure. Two inflection points are apparent. There is an initial period where there are a few failures (until about day 800), then a middle period of several failures (to about day 1 200), and finally a longer period of few failures, showing a slower rate of decline to the end of the study (Figure 12).

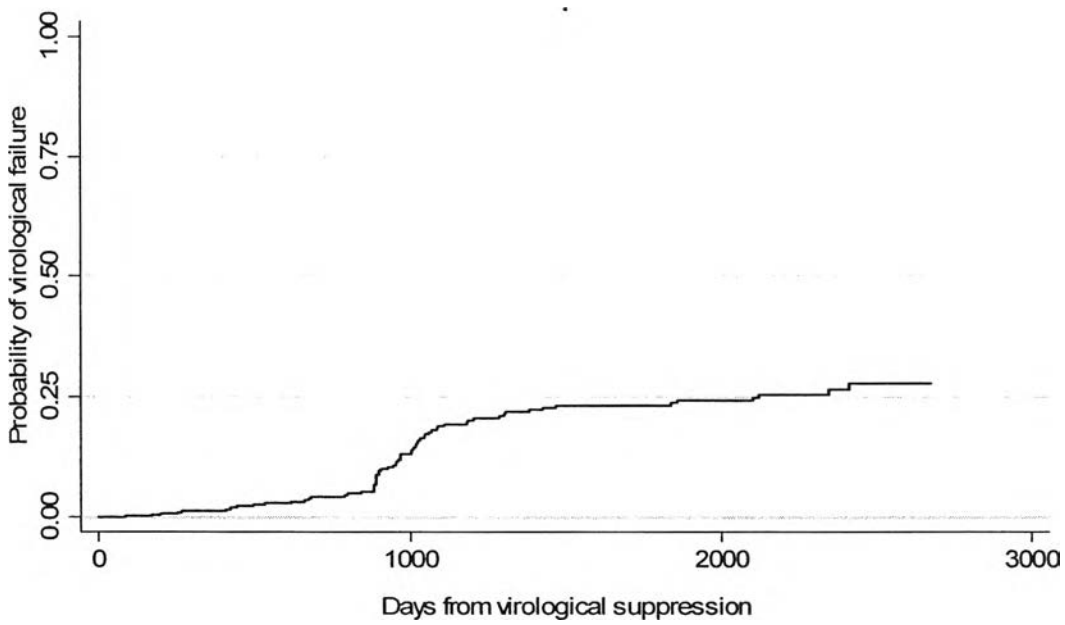


Figure 12: Kaplan-Meier curve for probability of virological failure

The curve in Figure 12 starts from zero, i.e., there are no failures at this point. These two inflection points suggested that polynomial transformation of the time variable from the expanded data set could be useful and might achieve a better fit in the logistic regression model, as in the period from start of study until suppression. Looking at the curve for the probability of virological failure, the two inflections

suggested quadratic and cubic transformation. This strategy was adopted for both the intervals in this study, from the start of HAART to suppression and from suppression to failure.

The variable for time achieved by the expansion of the data set was modelled by itself using logistic regression. The likelihood ratio χ^2 of the model was 0.5871, P=0.4435. When modelled alone including quadratic transformation of the time variable, the likelihood ratio χ^2 of the model was 33.2781, P<0.0001, indicating a better fit of the model. The addition of cubic transformation of the variable to the model produced an even better fit, the likelihood ratio χ^2 of the model was 44.6527, P<0.0001, demonstrating a steady increment in the model χ^2.

Bivariate Logistic Regression: from Suppression to Virological Failure

The time variable with quadratic and cubic transformation was modelled in bivariate analysis against 12 covariates: gender, baseline weight, mode of transmission (modelled as homosexual vs. heterosexual and other vs. heterosexual, and always modelled together), baseline CDC category, baseline age, year of entry, hepatitis B and C infection, CD4+ count, pVL, baseline regimen and whether or not they were suppressed by week 12 of study. The variables selected for multivariable analysis were the same as the analysis for the interval to suppression with the addition of the variable indicating whether or not participants were suppressed by week 12 after starting therapy. This variable proved to be a strong and significant predictor of failure (OR 2.18, P<0.0001).

Table 11: Results of bivariate logistic regression from suppression to virological failure

Independent variable time	**Likelihood ratio**	***P* value of the model**
Model with time and no other independent variables†	**44.6527**	**<0.0001**

†The time variable includes quadratic and cubic transformation

Independent variables		**OR**	**95% CL**	***P* value**
gender	female vs. male	0.613	(0.378 to 0.995)	0.0475*
weight at baseline	per kilogram	1.016	(0.996 to 1.037)	0.1204*
mode of transmission	homosexual vs. heterosexual	1.276	(0.750 to 2.171)	0.3686*
	other vs. heterosexual	0.323	(0.079 to 1.332)	0.1181*
baseline CDC category		0.881	(0.626 to 1,241)	0.4683
age at entry	in years	0.985	(0.952 to 1.020)	0.3911
year of entry	from 2000 to 2006	0.891	(0.729 to 1.088)	0.2559
Hep B status	positive vs. negative	0.901	(0.430 to 1.888)	0.7824
Hep C status	positive vs. negative	1.555	(0.672 to 3.595)	0.3021
CD4+ at baseline (cells/mm^3)	CD4+ ≥200 vs. <200	1.785	(1.103 to 2.890)	0.0184*
RNA at baseline	>50,000 copies/mL	1.591	(0.983 to 2.573)	0.0586*
baseline regimen	0=NNRTI, 1=PI	0.450	(0.215 to 0.941)	0.0338*
not suppressed by week 12		2.180	(1.142-4.160)	0.0181*

** selected for multivariable analysis*

OR = Odds ratio, CL = Wald Confidence Limits, LR = Likelihood Ratio χ^2 of the model

Table 11 above presents the results of bivariate logistic regression of the model from virological suppression to virological failure, including quadratic and cubic transformation of the time variable as previously described, and then the results of bivariate analysis that included the time variable with the same polynomial transformations modelled with each of the selected covariates.

These analyses resulted in eight variables ($p \leq 0.2$) being included in the multivariable analysis: *gender* [OR 0.613, 95% CL 0.378 to 0.995, *P*=0.0475],

These analyses resulted in eight variables ($p \leq 0.2$) being included in the multivariable analysis: *gender* [OR 0.613, 95% CL 0.378 to 0.995, *P*=0.0475], *baseline weight* [OR 1.016, 95% CL 0.996 to 1.037, *P*=0.1204], *mode of HIV transmission* (modelling together two variables, homosexual vs. heterosexual and other vs. heterosexual) [OR 1.276, 95% CL 0.750 to 2.171, *P*=0.3686 and OR 0.323, 95% CL 0.079 to 1.332, *P*=0.1181 respectively], *CD4+ count at baseline* (CD4+ ≥200 vs. CD4+ <200) [OR 1.785, 95% CL 1.103 to 2.890, *P*=0.0184], *pVL at baseline* (pVL >50,000) [OR 1.591, 95% CL 0.983 to 2.573, *P*=0.0586], *baseline ARV regimen* (NNRTI vs. PI) [OR 0.450, 95% CL 0.215 to 0.941, *P*=0.0338] and *not suppressed by week 12* [OR 2.180, 95% CL 1.142 to 4.160, *P*=0.0181].

Multiple logistic regression: from suppression to virological failure

The first step in multiple logistic regression was to model time with its polynomial transformations (as described previously) alone and then with the other selected covariates, to test the differences in the β estimates. All *P* values for the Wald χ^2 were <0.001. The variables remain robust, the Wald χ^2 remains highly significant, the β estimates do not change direction and their percent changes were small (Table 12).

Table 12: Results of β estimates for time alone and with covariates selected for multivariable analysis

Variable	β before	β after	% change
Days on study	0.0126	0.0125	0.8%
Days on study squared	-9.95E-6	-9.61E-6	3.4%
Days on study cubed	2.208E-9	2.118E-9	4.1%

The eight variables identified from the bivariate models then were modelled with the time variable and its quadratic and cubic polynomials using multiple logistic regression (Table 13).

In the multivariable model for time to virological failure, the χ^2 Likelihood Ratio of the model increased from 44.6527 to 80.5950 ($P<0.0001$), which means that the overall model χ^2 highly significantly increased by 35.9523 for the addition of the eight independent variables ($P<0.0001$).

Table 13: Results of multiple logistic regressions for time to virological failure including weight

		OR	(95% CL) *	*P* value
Days on study		1.013	(1.007 to 1.018)	<0.0001
Days on study squared		1.000	(1.000 to 1.000)	<0.0001
Days on study cubed		1.000	(1.000 to 1.000)	<0.0001
Gender	female vs. male	0.657	(0.340 to 1.269)	0.2111
Baseline weight	per kilogram	1.010	(0.985 to 1.037)	0.4335
Mode of transmission	homosexual vs. heterosexual	0.995	(0.539 to 1.834)	0.9859
	other vs. heterosexual	0.278	(0.064 to 1.206)	0.0873
Baseline CD4+ (cells/mm^3)	CD4+ ≥200 vs. <200	3.044	(1.778 to 5.212)	<0.0001
Baseline pVL	>50,000 copies/mL	1.937	(1.136 to 3.303)	0.0152
Baseline regimen	0=NNRTI, 1=PI	0.488	(0.222 to 1.073)	0.0742
Not suppressed by week 12		1.920	(0.974 to 3.785)	0.0596

Likelihood Ratio χ^2 of the model with 11 degrees of freedom is 80.5950 ($p<0.0001$)

* OR = Odds ratio; CL = Wald confidence limits

As described above, there was a highly statistically significant difference in the weights of males and females (t=11.22, P<0.0001). Weight is highly associated with gender and may thus confound the effect of gender in multiple regression models. After removing weight from the model, gender became statistically significant (P=0.0486) (Table 14). The χ^2 Likelihood Ratio of the model changed slightly to 79.9926 (P<0.0001), which means that the overall model χ^2 still highly significantly increased from the original model by 35.3399 for the addition of seven independent variables.

Table 14: Results of multiple logistic regressions for time to virological failure without weight

		OR	(95% CL) *	*P* value
Days on study		1.013	(1.007 to 1.018)	<0.0001
Days on study squared		1.000	(1.000 to 1.000)	<0.0001
Days on study cubed		1.000	(1.000 to 1.000)	0.0009
Gender	female vs. male	0.571	(0.328 to 0.996)	0.0486
Mode of transmission	homosexual vs. heterosexual	0.947	(0.519 to 1.727)	0.8583
	other vs. heterosexual	0.293	(0.068 to 1.266)	0.1003
Baseline CD4+ (cells/mm^3)	CD4+ ≥200 vs. <200	3.129	(1.836 to 5.331)	<0.0001
Baseline pVL	>50,000 copies/mL	1.936	(1.135 to 3.300)	0.0153
Baseline regimen	0=NNRTI, 1=PI	0.508	(0.233 to 1.108)	0.0889
Not suppressed by week 12		1.929	(0.977 to 3.809)	0.0584

Likelihood Ratio χ^2 of the model with 10 degrees of freedom is 79.9926 ($p<0.0001$)

* OR = Odds ratio; CL = Wald confidence limits

The final multivariable model of the time to virological failure identified three of the independent variables as statistically significant in predicting time to virological failure.

Both female gender (OR 0.571, 95% CL 0.328 to 0.996, *P*=0.0486) (Figure 13); and a baseline pVL ≥50,000 copies/mL of blood (OR 1.936, 95% CI 1.135 to 3.300, *P*=0.0135) were predictive of slower virological suppression. A baseline CD4+ count (OR 3.129, 95% CI 1.836 to 5.331, *P*<0.0001] (Figure 14) was predictive of faster virological suppression. This was an unexpected result and the reasons are not understood. Further analyses are needed to see if this unexpected result can be explained.

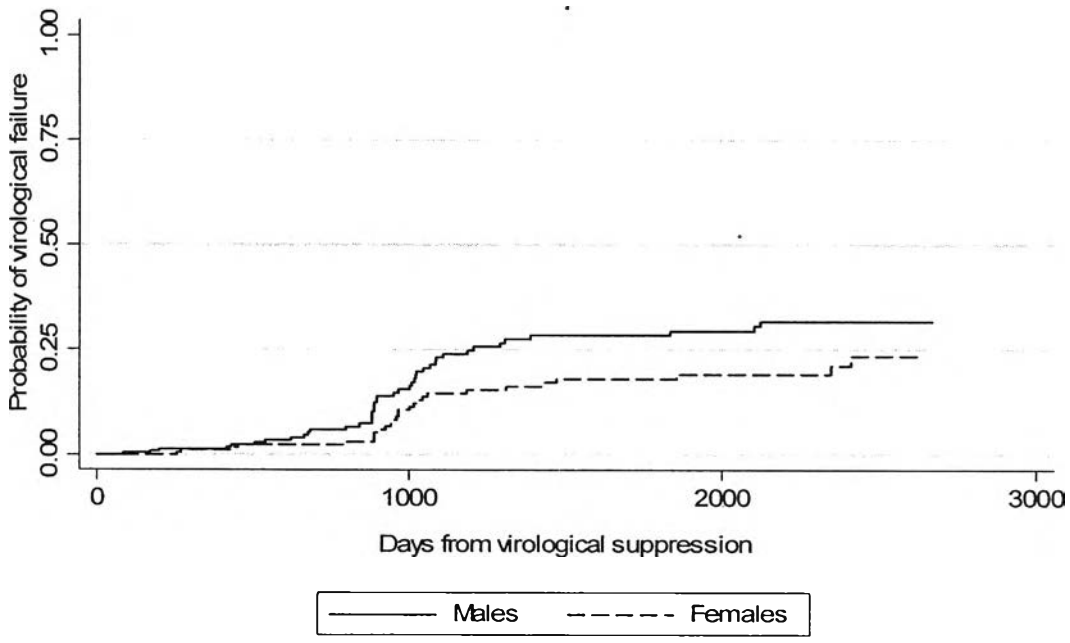


Figure 13: Kaplan-Meier curve for probability of virological failure stratified on gender

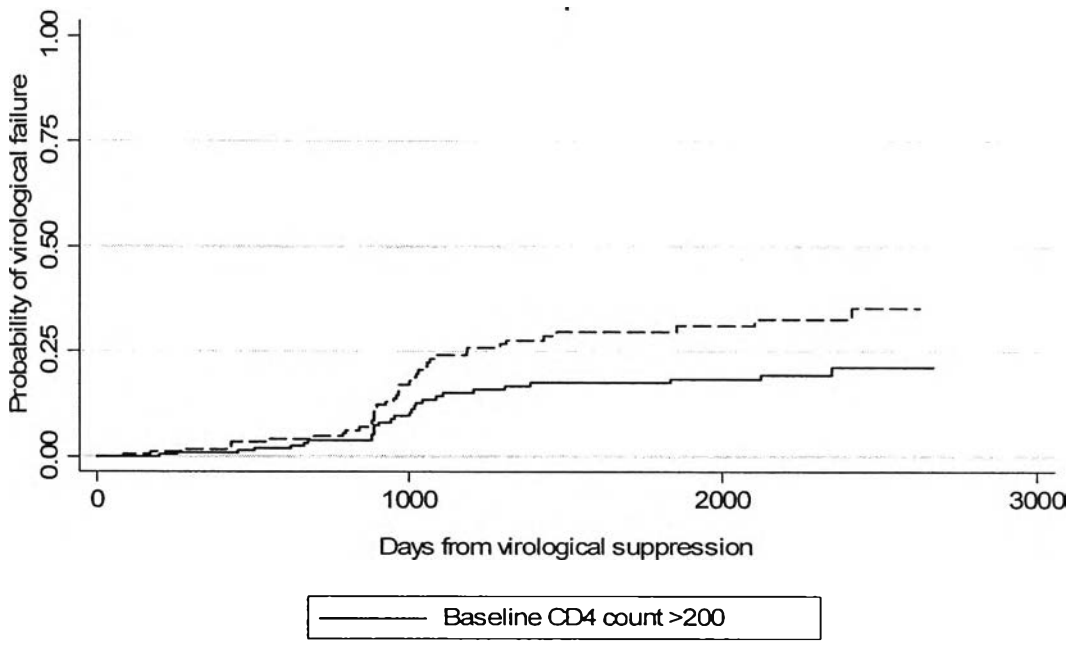


Figure 14: Kaplan-Meier curve for probability of virological failure stratified on CD4+ count at baseline

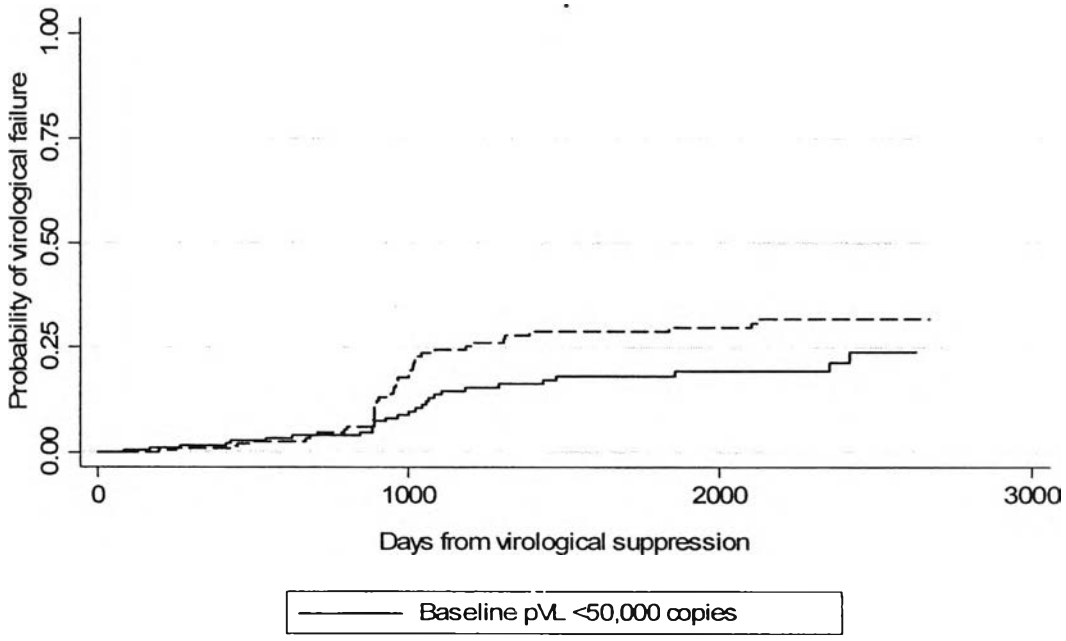


Figure 15: Kaplan-Meier curve for probability of virological failure stratified on plasma viral load

Additionally, two variables were not statistically significant but indicated a trend to predicting failure. A baseline regimen including a PI (OR 0.508, *P*=0.089) predicted slower virological failure and if not suppressed by week 12 (OR 1.929, *P*=0.058) predicted faster virological failure.

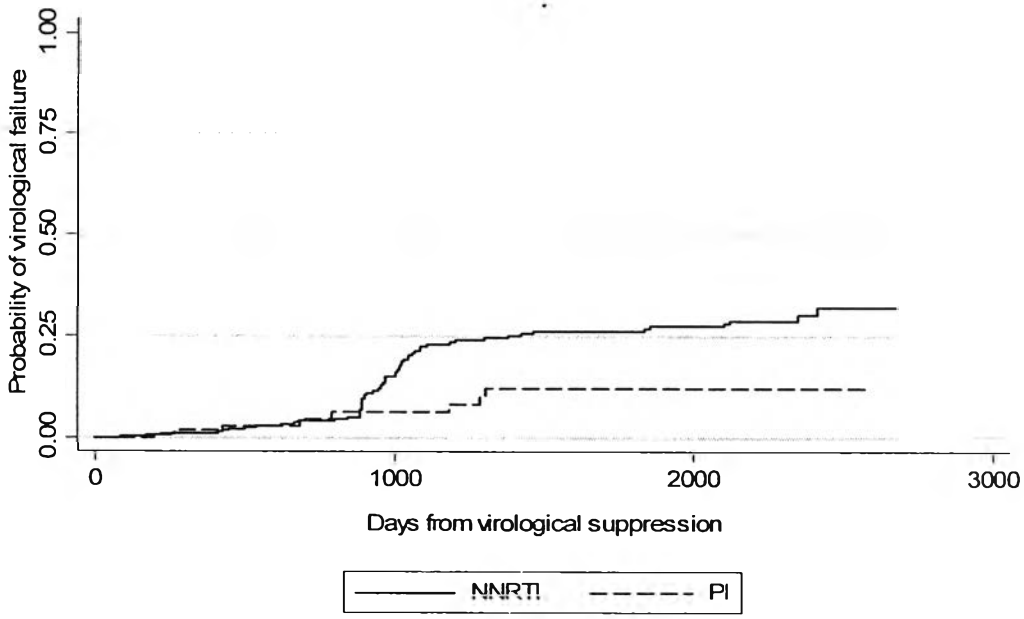


Figure 16: Kaplan-Meier curve for probability of virological failure stratified on baseline regimen

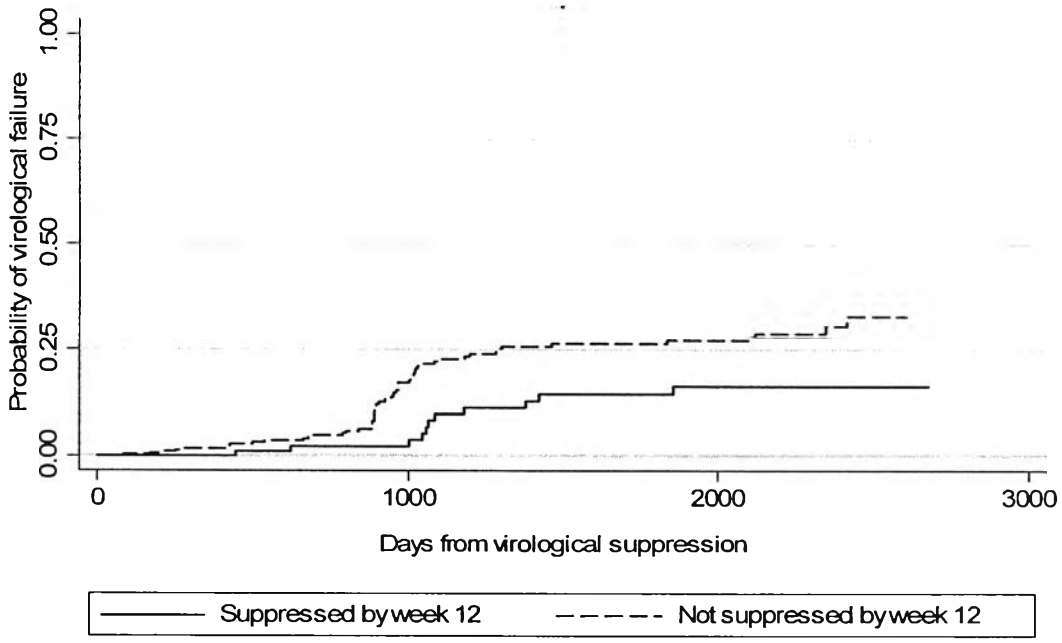


Figure 17: Kaplan-Meier curve for probability of virological failure stratified on time to virological suppression

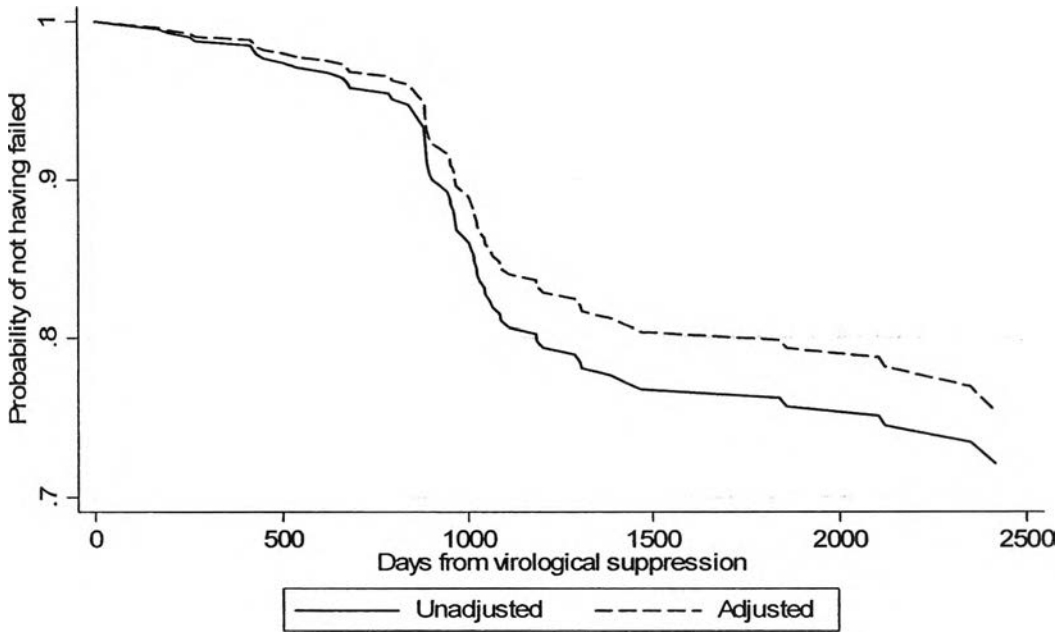
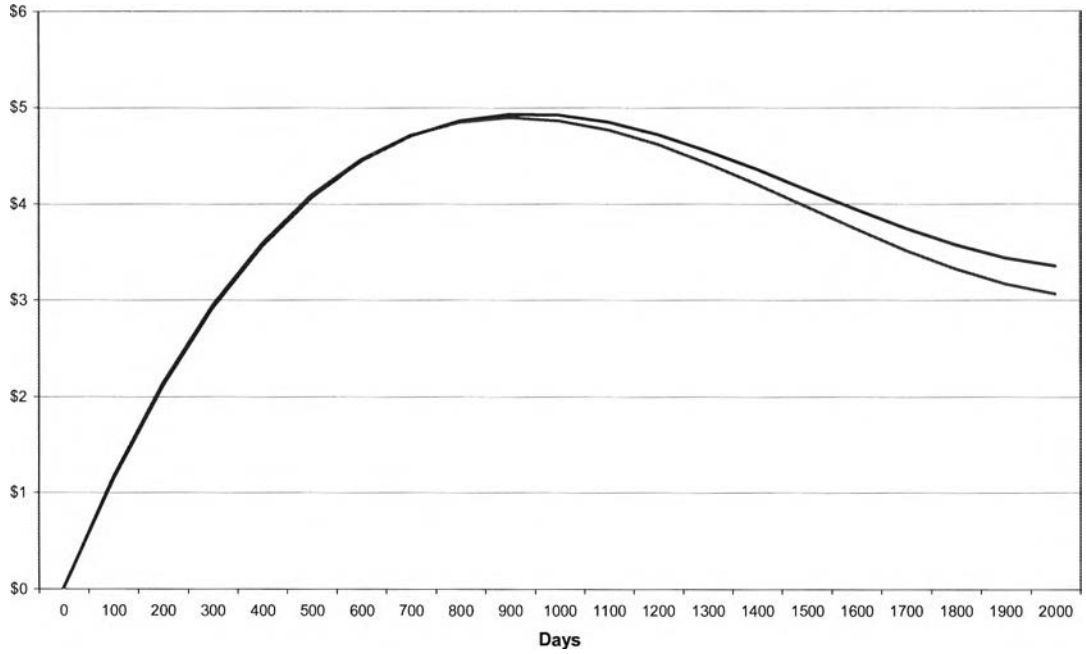


Figure 18: Overlay of survival curves for virological failure with the time variable unadjusted and adjusted

Figure 18 shows the overlay of the two survival curves for time to virological failure, starting with 100% probability of not having failed. It is clear to see that the shapes of the curves are very similar, showing that the polynomial transformation does not distort the shape. However, as demonstrated above, the likelihood ratio χ^2 of the transformed model provides a better fit. It is important to note that the scale of the y-axis in this table starts at 0.7 and not at zero. The apparent difference in these curves after day 1 000 is only about 1.6%.

Figure 19 shows the plot of the unadjusted and adjusted curves of the polynomial function of the time variable using log odds by days of follow-up for time from virological suppression to virological failure. These curves also illustrate the strength of the polynomial model and the stability of the time variable after polynomial transformation.



The green (lower) line is the unadjusted time variable, blue line the adjusted time variable

Figure19: Time from virological suppression to virological failure plotting log odds by days of follow-up