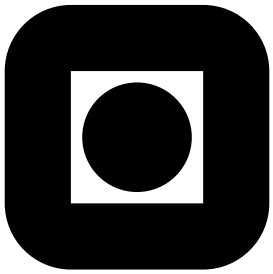
NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Statistical analysis of the Vacc-4x and Vacc-5q vaccine immunization studies for HIV-1 infected patients

by

Anastasia Ushakova

PREPRINT
STATISTICS NO. 2/2009



NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL http://www.math.ntnu.no/preprint/statistics/2009/S2-2009.ps
Anastasia Ushakova has homepage: http://www.math.ntnu.no/∼anastasi
E-mail: anastasi@stat.ntnu.no
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

Statistical analysis of the Vacc-4x and Vacc-5q vaccine immunization studies for HIV-1 infected patients

Anastasia Ushakova
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim, Norway

Abstract

In this article, we perform a statistical analysis of reoccurrence of opportunistic diseases for HIV-1 infected patients vaccinated with Vacc-4x and Vacc-5q therapeutic vaccines, comparing to a control group. The Peña-Hollander model for recurrent events is used to compare the vaccinated patients from the clinical studies and a control group of patients on HAART with similar conditions of the immune system.

Keywords: HIV-1, immunotherapy, recurrent events, therapeutic vaccines

1 Introduction

Highly active antiretroviral therapy (HAART) leads to a prolonged reduction of viral load in many patients. However, HIV-1 specific immune responses are reported to be declined under continuous therapy (probably, because of the drop of viral load), while T-cell responsiveness to other infections are restored during suppressed viremia. Immunotherapy is aimed to reactivate the HIV-specific immune responses. This could be achieved by vaccinating the patient with a special vaccine, or potentially by therapy interruption.

p24 is the core protein in the HIV virus particle forming the capsid of the virus. HIV-specific immune responses are directed against different parts of the virus, and responses against p24 have proved to be of particular importance in controlling HIV-replication.

The Vacc-4x immunotherapy candidate is a vaccine composed of four modified water-soluble peptides (20-27 amino acids in length) corresponding to conserved domains of the HIV-1 protein p24 that preferentially include HLA-A2 restricted elements. The vaccine has been described in detail in Åsjö et al. (2002). The Vacc-5q peptide-based pentavalent therapeutic vaccine consists of short consensus peptides from p17, p24, and Tat.

The Vacc-4x study was organized as follows: 40 patients stable on HAART, with a median value of 550 CD4+ T cells/μl and HIV RNA $<$ 50 copies/ml at inclusion, were randomized into two dosage-arms receiving 0.4 and 1.2 mg Vacc-4x peptides, respectively. In the first phase of the study, described in Kran et al. (2004), the vaccine was injected intradermally over a period of 26 weeks under continuous HAART. In the second phase of the study, described in Kran et al. (2005), Kran et al. (2006), the HAART was interrupted two times: first for 4 weeks, then for 14 weeks, with the period

under HAART lasting for 8 weeks in between. See Figure 1 for the schematization of the study. The Vacc-5q study design and immunization schedule were similar and 10 patients were involved. After week 52, in both studies, 11 patients recommenced continuous HAART (7 patients from Vacc-4x and 4 patients from Vacc-5q), while the remaining 36 patients did not.

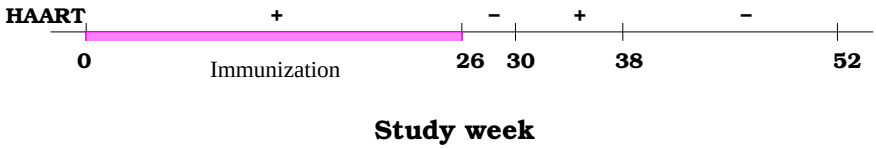


Figure 1: The schedule of the Vacc-4x clinical study.

In Kran et al. (2005), Kran et al. (2006) the clinical efficiency of the vaccination with Vacc-4x in phase II of the study was statistically analyzed in terms of delayed-type hypersensitivity reaction (DTH) and the ability to achieve lower actual HIV RNA levels. In this work the analysis of the vaccinated patients is continued. Of possible interest is to investigate the probable impact of the immunization on the general characteristics of the patients from the clinical study group, and to compare those with the control group. All the data come from the Ullevaal Hospital database for HIV-infected patients in Oslo.

The statistical analysis is carried out using R (R Development Core Team, 2008). To fit the Pena-Hollander model gcmrec package (González et al., 2008) is used.

The paper is organized as follows. In Section 2 we define the study groups and investigate if these groups are comparable. In Section 3 we compare the distributions of the time to occurrence of first event after the subject enters the study for these three groups. In Section 4 we introduce the Peña-Hollander model for recurrent events and compare the study groups with respect to their distributions of event reoccurrences. We conclude with a summary in Section 5.

2 Clinical efficacy outcome

One of the indicators of the immune system's strength is the frequency of secondary diseases that the patient experiences being HIV-positive. In the database we have the longitudinal observations of occurrence of the opportunistic diseases listed in Table 4 in Appendix. We call an occurrence of a disease from this list an *event*. On the one hand, the events from this group only, may not describe good enough the immune system's impairment, since they do not contain all the sicknesses that the patients have experienced. On the other hand, since the same sicknesses reoccurrences are available for patients from all the groups, they definitely have descriptive properties.

Among a total of 47 vaccinated patients, 37 patients belonged to the Vacc-4x trial, while the other 10 belonged to the Vacc-5q trial. Thus we split these patients into two groups, G_1 and G_2, containing the 37 and 10 patients respectively.

We compare the distribution of the time to events for the groups G_1, G_2 and an additional group of control patients with approximately the same CD4+ T cells counts, being on therapy (containing all the individuals that are available from the Ullevaal database). For the groups G_1 and G_2, the reference

point (time 0) is the time when the patient was first injected with Vacc-4c and Vacc-5q, respectively; for the control group, the reference time is the first time on therapy when the patient had CD4+ T cells count $\in (346.3, 890.0)$, i.e. between the first and third quartile of the CD4+ T cells counts for the union of the groups G_1 and G_2 at the reference point. However, a large part of the patients selected like this appear to have been diagnosed with HIV infection quite early and started their therapy in the eighties or early nineties, and therefore, did not have an opportunity to take multiple-drug therapy, which could had influenced their health conditions resulting in a higher occurrence of events. So we have narrowed the group of controls, restricting it to only those individuals that had received multiple therapy (at least 3 different drugs) and that started the first drug from the list of therapy (mostly AZT or 3TC) not earlier than 01.01.1998. The control group consists of 298 patients.

2.1 CD4 count distribution at the event occurrence

In order to justify that these three groups are fairly comparable, we compare their distributions of CD4 count at the time of every event's occurrence after the reference point, using the Kolmogorov-Smirnov test.

H_0	H_1	p-value	conclusion
$F_{G_1}(t) = F_{G_2}(t)$	$\neq$	0.821	H_0 accepted
$F_{\{G_1 \cup G_2\}}(t) = F_{\text{control}}(t)$	$\neq$	0.3145	H_0 accepted

2.2 The number of events for the three groups before the reference time

One may argue that the criterion for inclusion to the study, based on CD4+ T cells count may not be good enough since the immune impairment under the HIV infection is a too complicated process which cannot be exhaustively classified by just this measurement. For example, it might be argued that our analysis could be influenced by a possible choice of more healthy individuals for the clinical study compared to the control group.

Thus we compare the groups with respect to the distribution of the number of events that had occurred prior to the reference time. Histograms of these distributions are shown in Figure 2. The estimated means and medians are

	G_1	G_2	Control
mean	4.189189	5.300000	2.845638
median	3	5	2

The Kolmogorov-Smirnov test gives the following results:

H_0	H_1	p-value	conclusion
$F_{G_1}(t) = F_{G_2}(t)$	$\neq$	0.2379	H_0 accepted
$F_{\{G_1 \cup G_2\}}(t) = F_{\text{control}}(t)$	$\leq$	0.001004	H_0 rejected

The conclusion is thus that the number of initial events in the group $\{G_1 \cup G_2\}$ is stochastically larger than in the control group, which even improves the evidence of positive influence of the vaccination, if any found.

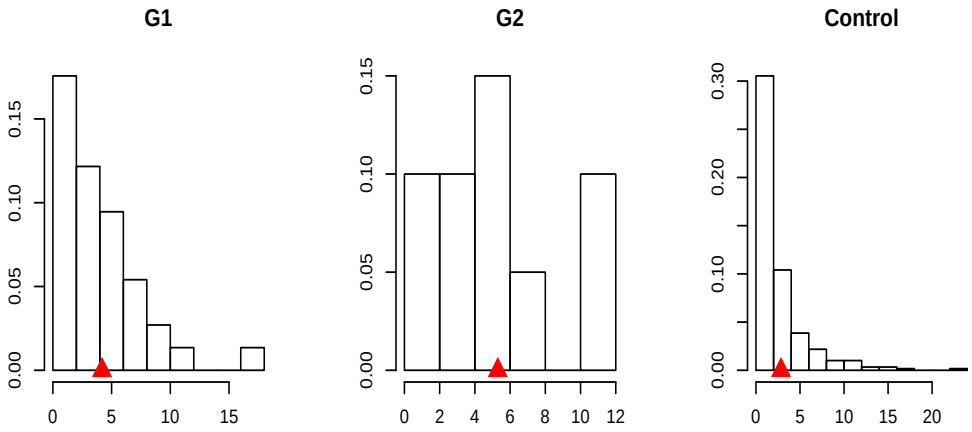


Figure 2: Histograms of the distributions of events prior to the entering to the study (densities). X-axis show the number of events.

3 Time until the first event

The Kaplan-Meier plots of the survival function of the first event after the reference point are plotted in Figure 3, and the results of the log-rank test for the survival function of the first event are presented in Table 1. For this choice of the control group, there is no significant difference in the time until the

Groups compared	p-value
G_1 vs Controls	0.297
G_2 vs Controls	0.3
G_1 vs G_2	0.0595
$\{G_1 \cup G_2\}$ vs Controls	0.582

Table 1: Results of the log-rank test for survival function of the first event for the control group.

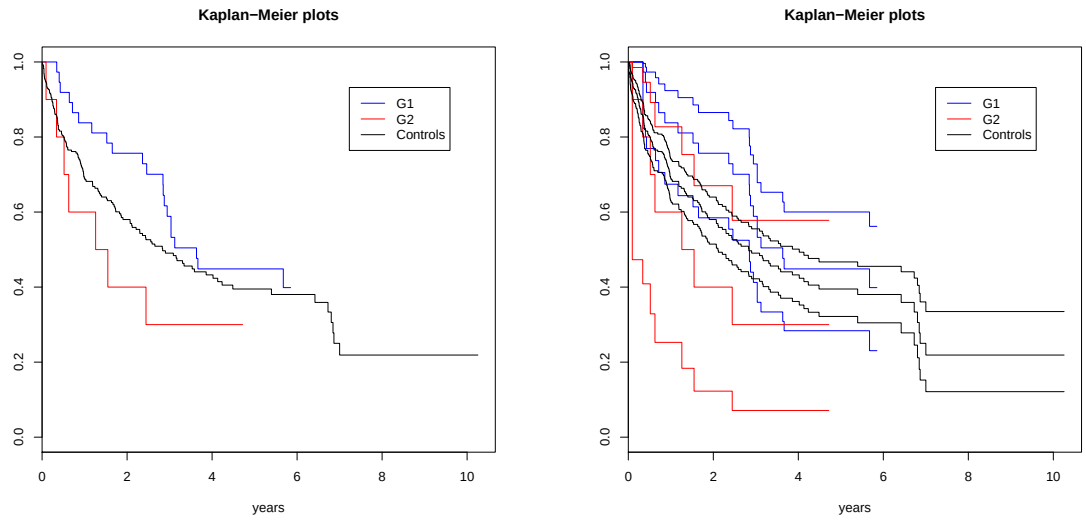
first event for both of the vaccinated groups, however, a close to 5%-significant difference between G_1 and G_2.

For a more detailed analysis we will next utilize the times for all the events listed.

4 Analysis of reoccurrence of events

In this section we use the semiparametric model for recurrent event data proposed by Peña and Hollander (2004)[1].

[1]Program language R, package 'gcmrec' was used for the analysis.



(a) (b)

Figure 3: Kaplan-Meier plots of distribution until the first event with the group of controls, (a) without confidence intervals and (b) with estimated 95% confidence intervals.

4.1 Basic model

Let us define for patient i, the vector of covariates $X_i(t) = (X_{1i}(t), X_{2i}(t), \ldots, X_{ik}(t))$. The model we use has intensity function of the form

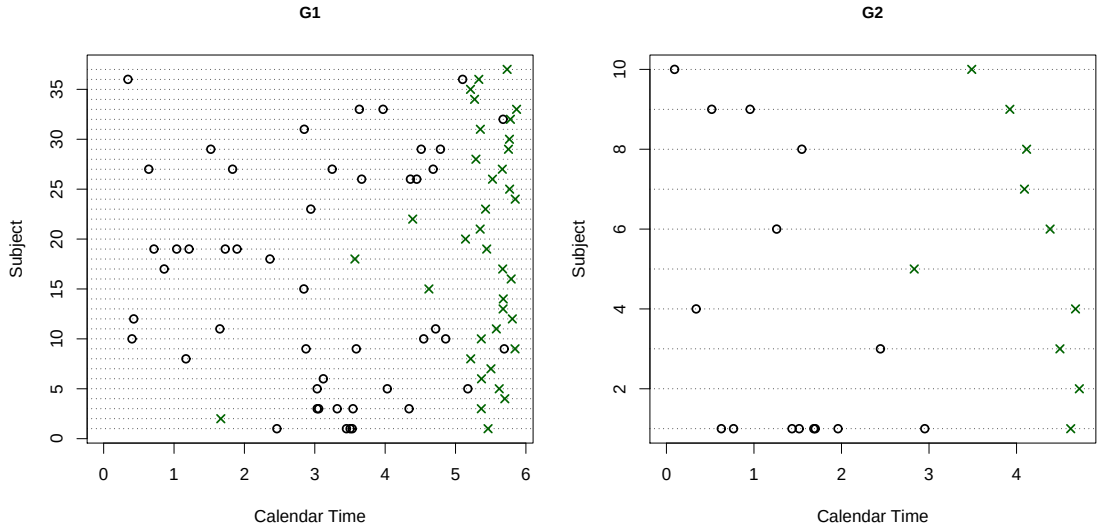
$$\lambda_i(t|X_i(t)) = \lambda_0(t)\alpha^{N_i(t)} \exp\{\beta^T X_i(t)\}, \quad t \geq 0$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, $N_i(t)$ is the number of events before time t (only those events that occurred after the reference point are counted) and $\beta = \{\beta_1, \beta_2, \ldots, \beta_k\}^T$ is the regression parameter (see paper Peña et al. (2007) for details). The term $\alpha^{N_i(t)}$ is attributable to the accumulating event occurrences. Thus, $\alpha > 1$ models a situation where an increasing number of event occurrences has an adverse effect, $\alpha = 1$ – no effect, and $\alpha < 1$ – beneficial effect.

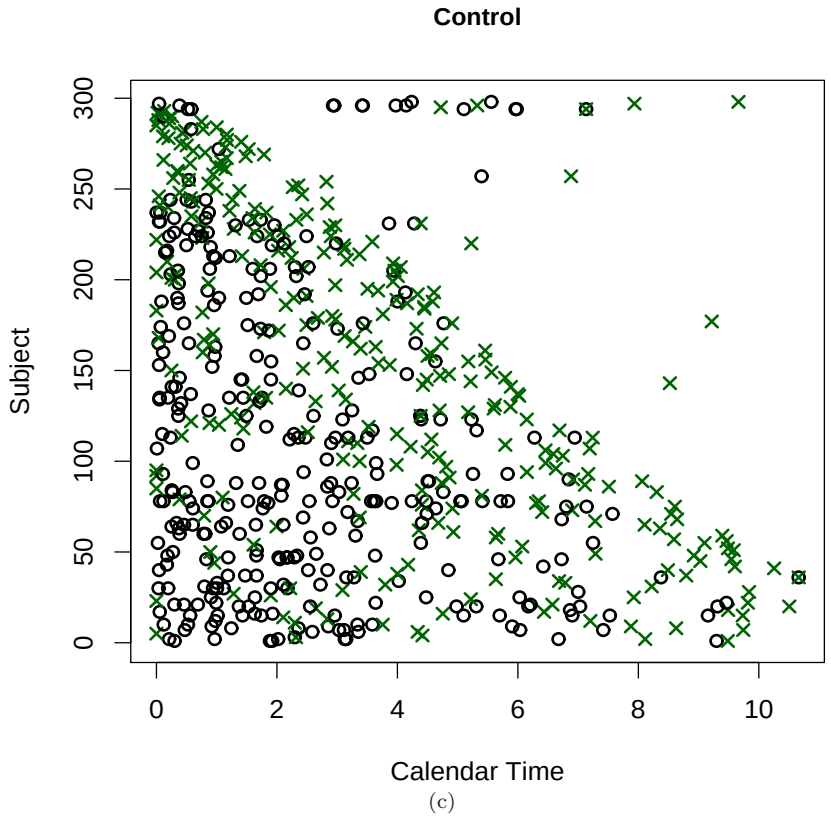
The data of occurrence of events is summarized in Figure 4. For every patient, there are either observed events (marked on the graph with the symbol '`o`'), or right-censored observations (marked with '`x`'), when it is known that no event had occurred during this time, but the patient drops out of the study, i.e. no information is available about him or her after this time point. The groups G_1, G_2 and the control group include 36, 11 and 368 patients respectively.

We define the covariates as follows. Covariate X_1 relates to the group G_2: $\{X_1 = 1\}$ if the observation comes from the patient belonging to group G_1, and $\{X_1 = 0\}$ otherwise. Covariate $\{X_2 = 1\}$ if the patient belongs to the control group, and $\{X_2 = 0\}$ otherwise. Covariate X_3 is the number of events that already occurred before the time that the patient was included to the study.

Statistical results of the estimation are presented in Table **??**. Estimated baseline survival function corresponding to the statistical results given in Table **??** is presented in Figure 5.



(a) (b)



(c)

Figure 4: Plots of successive occurrences of events for the groups: G_1 (a), G_2 (b) and the control group (c). The time scale is years.

Term	Parameter	coef	exp{coef} (sd)	P-value
G_2 group indicator	β_1	0.0554	1.06 (0.2990)	0.85
Control group indicator	β_2	0.2488	1.28 (0.1590)	0.12
Number of events at the reference point	β_3	0.0981	1.10 (0.0123)	$1.7 \cdot 10^{-15}$
Impact of accumulating events	α	1.17 (0.0257)		
Log-likelihood		-2019.45		

Table 2: Basic model.

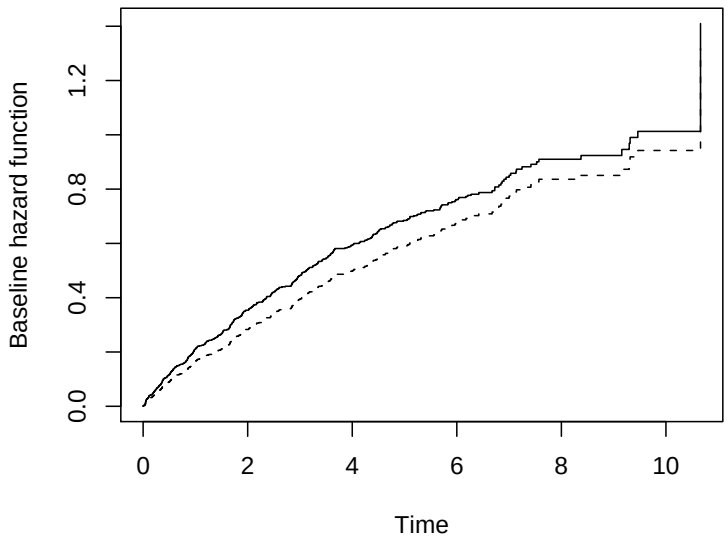


Figure 5: Estimated baseline hazard function for the basic model (solid line) and model with frailty (broken line).

4.2 Model with frailty component

For the model of Peña and Hollander, frailties are defined as a vector of independent identically distributed random variables, $Z = (Z_1, Z_2, \ldots, Z_n)$ from a parametric distribution $H(z;\xi) = P(Z \leq z|\xi)$ where ξ is a finite-dimensional parameter taking values in $\Xi \subseteq \mathbb{R}^r$. These variables are unobservable random factors that influence the event reoccurrences. The conditional intensity function for the model with frailties has the form

$$\lambda_i(t|Z_i, X_i(t)) = Z_i \lambda_0(t) \alpha^{N_i(t)} \exp\{\beta^T X_i(t)\}, \quad t \geq 0$$

The distribution $H(z;\xi)$ is chosen to be the gamma distribution with unit mean and variance $1/\xi$, $H = Gamma(\xi, \xi)$ and parameter ξ is to be estimated. The results of the estimation are presented in Table 3. For this model, the estimated $\hat{\beta}_2$ is significantly greater than zero. The estimated likelihood

Term	Parameter	coef	exp{coef} (sd)	P-value
G_2 group indicator	β_1	0.106	1.11 (0.4596)	0.82
Control group indicator	β_2	0.383	1.47 (0.1994)	0.054
Number of events at the reference point	β_3	0.135	1.14 (0.0212)	$2 \cdot 10^{-10}$
Impact of accumulating events	α	1.03 (0.0481)		
Frailty	ξ	2.26 (0.805)		
Log-likelihood		-2035.77		

Table 3: Model with frailty component.

for the model with frailty component is lower than the likelihood for the model without frailty due to the difference of methods of the likelihood evaluation for models with and without frailty component, see Peña et al. (2007) for the details.

Thus, the clinical efficiency of vaccination with Vacc-4x and Vacc-5q is profitable comparing with the control group with p-value close to 5%, while there is no statistical difference between the two vaccinated groups. The estimated frailty is statistically significant, which implies that there could be found other covariates that significantly influence the hazard function.

5 Conclusion

A statistical analysis shows a close to significant at 5% level profit of therapeutic vaccines Vacc-4x and Vacc-5q combined with treatment interruptions for the patient's immune system measured in terms of probability of reoccurrence of adverse events, compared to a reference group of patients with similar CD4 T cells count at the entrance of the study, and better history of adverse events at the entrance. There was also no significant difference for the vaccinated individuals in the occurrence of adverse events between the groups vaccinated with Vacc-4x and Vacc-5q therapeutic vaccines. We suggest that vaccination with therapeutic vaccines Vacc-4x and Vacc-5q is a promising approach and it's effect should be investigated further.

Appendix

The list of opportunistic diseases whose occurrence is defined as an event is presented in Table 4.

Code	**Event**	**Code**	**Event**
C100	Kidney stones	C68	Acute cytomegalovirus (CMV) infection
C101	Diabetes mellitus	C69	Otitis
C102	Pancreatitis	C70	Varicella
C103	Lipodystrophy	C71	Bronchitis
C104	Thrombosis	C72	Tonsillitis/Pharyngitis

C105	Neuropathy	C73	Gingivitis
C106	Verruca vulgaris	C74	Parvovirus infection
C107	Arteriosclerosis obliterans	C75	Condyloma
C108	Immune reconstitution inflammatory syndrome (IRIS)	C76	Syphilis (Lues)
C109	Hypertension	C77	Chlamydia infection
C47	Sinusitis	C78	Gonorrhoe
C48	Abscess/Phlegmon	C79	Vaginal candidiasis
C49	Phlegmon	C80	Salmonellose
C50	Pyoderma/Skin infection	C81	Amoebiasis
C51	Urinary tract infection	C82	Shigellosis
C52	Herpes labialis < 1 mnd.	C83	Giardiasis
C53	Herpes genitalis < 1 mnd.	C84	Clostridium difficile infection
C54	HSV infection, unspecified	C85	Diarrhea non-ulcer dyspepsia
C55	Staphylococcus infection	C86	Appendicitis
C56	Streptococcus infection	C87	Salpingitis
C57	Haemophilus influenzae infection	C88	Meningitis
C58	Pneumococcus infection	C89	Diarrhea campylobacter
C59	Sepsis	C90	Hepatitis A acute
C60	Dermatitis	C91	Hepatitis B acute
C61	Psoriasis	C92	Hepatitis C acute
C62	Rhodococcus sepsis	C93	Hepatitis NUD acute
C63	Arthritis	C95	Myocardial infarction
C64	Molluscum contagiosum	C96	Mycobacterial infection local
C65	Pityriasis rosea	C97	Cancer
C67	Mononucleosis	C99	Heart disease

Table 4: List of events.

References

[1] González, J. R., Slate E. H., Peña, E. A. (2008) gcmrec: General class of models for recurrent event data. R package version 1.0-2. http://www.r-project.org

[2] Kran, A.-M. B., Sørensen, B., Nyhus, J., Sommerfelt, M. A., Baksaas, I., Bruun, J. N., Kvale, D. (2004) HLA- and dose-dependent immunogenicity of a peptide-based HIV-1 immunotherapy candidate (Vacc-4x). *AIDS* **18** (14), 1875–1883.

[3] Åsjö, B., Stavang, H., Sørensen, B., Baksaas, I., Nyhus, J., Langeland, N. (2002) Phase I trial of a therapeutic HIV type 1 vaccine, Vacc-4x, in HIV type 1-infected individuals with or without antiretroviral therapy. *AIDS Res. Hum. Retroviruses* **20**, 621–667.

[4] Kran, A.-M. B., Sommerfelt, M. A., Sørensen, B., Nyhus, J., Baksaas, I., Kvale, D. (2005) Reduced viral burden amongst high responder patients following HIV-1 p24 peptide-based therapeutic immunization. *Vaccine* **23**, 4011–4015.

[5] Kran, A.-M. B., Sørensen, B., Sommerfelt, M. A., Nyhus, J., Baksaas, I., Kvale, D. (2006) Long-term HIV-specific responses and delayed resumption of antiretroviral therapy after peptide immunization targeting dendritic cells. *AIDS* **20**, 627–630.

[6] Kvale, D., Aaberge, I. S., Frøland, S. S. (2002) Predictive value of immunologic parameters and HIV RNA in relation to Humoral pneumococcal polysaccharide vaccine responses in patients with HIV infection. *Eur. J. Clin. Microbiol. Infect. Dis.* **21**, 688–690.

[7] Peña, E. A., Slate, E. H., Gonzales, J. R. (2007) Semiparametric inference for a general class of models for recurrent events. *J. of Statistical planning and Inference*, **37**, 1727–1747.

[8] R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.