Assessment of the Relative Therapeutic Effect in Small Groups at Several Time Points: Comparison of Mucosal and Subcutaneous Peptide Vaccines in Rhesus macaques Exposed to SHIV

Kuznetsov VA (¶,#), Stepanov VS(+), Berzofsky JA(*), Belyakov IM(*)

¶ *SRA International, Inc. 12a South Drive, Rm. 1018, NIH campus, Bethesda, MD, 20892-5772, USA; + Central Economic and Mathematics Institute, Russian Acad. of Sci., Moscow, 117418; *Molecular Immunogenetics and Vaccine Research Section, Vaccine Branch, National Cancer Institute, Building 10, 6B-12, Bethesda, MD 20892, USA; # correspondence should be addressed to V.A.K.: vk28u@nih.gov*

Abstract

Background: **Due to high cost, subject availability and ethical constraints, it is often critically important in pre-clinical and clinical studies to carry out an adequate statistical analysis of longitudinal multivariate data over several time-points in trials in several small groups.** ***Objectives:*** **We aim to accurately assess and develop an appropriate distribution-free longitudinal model for an estimate of the comparative treatment effects of several biological factors in several small groups even if the data sets should contain outlier measurements and censored values. This approach is used to evaluate the relative efficacy of mucosal and subcutaneous polypeptide vaccines in rhesus macaques exposed to SHIV.**

Study design: **The algorithms of the non-parametric repeated measures ANOVA models are described, programmed and assessed.**

Biological Results: **Using nonparametric ANOVA tests, we provided a statistical evaluation of the relative efficacy of mucosal and subcutaneous synthetic HIV/SIV peptide vaccines in rhesus macaques mucosally exposed to pathogenic SHIV-Ku2. We demonstrated with statistical significance that during the chronic phase of mucosal SHIV-Ku2 infection in immunized macaques, the numbers of CD4(+) and CD8(+) cells reciprocally reflect the virus titers in plasma, and both immune markers demonstrate better protection against pathogenic SHIV-Ku2 in intrarectally immunized MamuA*01 macaques.**

Conclusion: **Despite limited data, our analysis shows that there is a better preservation of both CD4(+) and CD8(+) cells in intrarectally immunized animals. Our analytical methodology can be applied in comparative estimates of the different treatment-associated effects and their synergy for a variety of longitudinal data sets (laboratory data, microarray data, clinical information), in small treatment groups, even if data sets contain outlier measurements and censored values.**

Keywords*:* relative treatment effect, non-parametric repeated measures ANOVA, longitudinal data analysis, small groups, SHIV, mucosal and subcutaneous vaccines, pre-clinical and clinical trials

1. INTRODUCTION

It is well known that high cost, subject availability and ethical constraints lead to the non-optimal design of many non-human primate studies and initial-phase clinical trials. In this situation, there is a high risk of failure. To overcome the problem of limited numbers of subjects, when they are studied longitudinally over time, we developed a novel approach using a non-parametric repeated measures ANOVA, based on the works of Brunner, Puri, and Akritas (BAP) (Akritas, Brunner, 1997; Brunner et al, 1999; Brunner, Puri, 2001) as well as the work of Mack (Mack, 1981).

In pre-clinical studies, a limited number Rhesus Macaques with a well characterized MHC class I molecule – MamuA*01 are typically used as an appropriate experimental animal model to evaluate prospective vaccines against HIV. However, because of the limited number of Rhesus Macaques per experimental group and their high cost, it is hard to obtain statistically reliable results. It is especially difficult to evaluate the relative therapeutic effects of vaccination on viral load in plasma and preservation of the number and functions of CD4(+) and CD8(+) cells, which are used as major markers of protection against retroviral challenge in pre-clinical and clinical trails. In particularly, in evaluating the CD8(+) and CD4(+) cell dynamics and SHIV virus load after fast acute phase of the SHIV infection at and over the steady-state (chronic) stage of the infection process we face the hurdle that, at this stage of infection, the virus load and the number of CD4(+) cells in blood show slow temporal trends for a long period of time and often have very similar dynamics for poor and good responders, with remarkable variation of these biological markers around their median values. However, careful analysis that considers the virus load in the blood and regional lymph tissues, the numbers of CD4(+) and CD8(+) cells in the blood, and the local mucosal SIV-specific CD8(+) CTL immune responses allows one to determine the significant correlates of these markers with the treatment outcome after SHIV-Ku2 mucosal challenge (Belyakov et al., 2001).

Here, we present the detailed algorithms of several modern non-parametric repeated measurements ANOVA tests. We aim to accurately assess and develop an appropriate distribution-free longitudinal model for an estimate of the comparative treatment effects of several biological factors in several small groups even when data sets contain outlier measurements and censored values. Using the tests we would like to answer three major questions:

(1) Are there different treatment effects, induced by intrarectal (IR) vs subcutaneous immunization of rhesus macaques with a synthetic SIV/HIV multi-epitope peptide vaccine in the presence of LT(R192G) mucosal adjuvant, on the numbers of virus particles in the blood and in the gut intestine, and the number of CD4(+) cells and CD8(+) cells in the blood of monkeys at early time points during chronic SHIVKu2 infection?

(2) Which method of vaccination (mucosal vs systemic) is more effective for long-term protection?

(3) Which method shows a significant improvement first and how long does that improvement last?

To answer to these questions, four null hypotheses were tested on our data:

(i) no group treatment effect (no statistical differences between levels of viremia) associated with different methods of vaccination,
(ii) no time-related treatment effect (no differences in the shape of kinetic curves of viremia decay)
(iii) no interaction synergism between group treatment effect and time-related treatment effect, and
(iv) no simple time effect within the groups.

In this work, using Mack's and BAP models, we develop the algorithms for statistical evaluation of the relative efficacy of mucosal and subcutaneous HIV/SIV vaccines in rhesus macaques mucosally exposed to pathogenic SHIV-Ku2. We show that during the chronic phase of mucosal SHIV-Ku2 infection in immunized macaques, the numbers of the CD4(+) and CD8(+) cells reciprocally reflect the virus titers in plasma and that both parameters demonstrate that there is better protection against pathogenic SHIV-Ku2 in intrarectally immunized MamuA*01 macaques.

2. METHODS

2.1 Animals, materials and experimental methods

2.1.1. Macaques and vaccines

The experimental design and partially resulting data were described previously in detail (Belyakov et al., 2001). The 11 Indian rhesus macaques (Macaca mulatta) were Mamu-A*01-positive for MHC class I.

2.1.2. Kinetic observation

The viral load was measured in the plasma of each monkey using stadard NASBA method on day 3, 10, 14, 21, 35, 42, 49, 56, 63, 70, 84, 120, 150 and 196 after challenge with SHIV-Ku2. The CD4(+) and CD8(+) cell counts in the blood were measured 15 days before SHIV-Ku2 virus challenge and on day 0, 10, 21, 35, 99, 120, 149 and 176 after challenge with SHIV-Ku2 (Table 1). The proliferative response of peripheral blood lymphocytes against PHA was measured on 90 day after rhesus macaques exposed to SHIV-Ku2.

Table 1. Design of kinetic experiments for the CD4, CD8 counts and virus load. Day 0 is the SHIV exposed day.

Table 1

Day	-15	0	3	10	14	21	35	42	49	56	63	70	84	99	120	149	150	176	196
CD4, CD8	**x**	**x**		**x**		**x**	**x**							**x**	**x**	**x**		**x**	
RNA			**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**		**x**		**x**		**x**

2.2. Longitudinal Non-parametric ANOVA-Type Models

2.2.1. The Brunner-Akritas-Puri (BAP) model

We consider the two and more groups comparisons. Let a denote the number of the groups. Let $A, B, C, ...$ denote the groups of interest numerated as the following $i = 1,...,a$. If these groups each consisting of n_i subjects are repeatedly observed under t different time points, then the repeated measures design, called the split-plot design, is an appropriate model the data of the experiment. In this model, $i = 1,...,a$ groups (factor G) each consisting of $k = 1,...,n_i$ subjects are observed on $j = 1,...,t$ occasions. In this study, $a = 2,3$, $k = 2,3,4$ and $t = 4,5,6,7,8,9,10$. Let n denote the total number of subjects in the experiment, then $n = \sum_{i=1}^{a} n_i$. Let $N = n \cdot t$ denote the number of observations (if there are no missing values).

The statistical model for t longitudinal observations of k-th subjects in group i can be written as the independent random vector

$$\mathbf{X_{ik}} = (X_{ik1},..., X_{ikt})', \quad i = 1,...,a, \quad k = 1,...,n_i \ , \tag{1}$$

which denotes the observation vector (i.e. the number of virus particles in the blood) of subject k within group i. These n vectors (i.e. subjects) are assumed to be independent. The components of each one of these vectors, however, can be arbitrarily dependent on one an another. Since the observations of different subjects within one group are considered as replications of the experiments, it is reasonable to assume that the distribution functions of vectors $\mathbf{X_{ik}} = (X_{ik1},..., X_{ikt})'$ do not depend on the index k. Within this framework, each X_{ikj}, $j = 1,...,t$, will have the distribution function $F_{ij}(x)$, where

$$F_{ij}(x) = 0.5 \cdot [F_{ij}^{+}(x) + F_{ij}^{-}(x)] \tag{2}$$

denotes the normalized version of the distribution function of random variable $\boldsymbol{X}$, where $F_{ij}^{+}(x) = P(X_{ikj} \le x)$ and $F_{ij}^{-}(x) = P(X_{ikj} < x)$ (Akritas, Brunner,1997).

The next four statistical hypotheses versus alternative hypotheses are considered:

1. Let $\bar{F}_{i\bullet} := E_i(F) = (\sum_{j=1}^{t} F_{ij})/t$ denote the average values of the distribution F_{ij} over all time points for the i-th group. The following hypothesis of the no treatment effect (group-associated factor) differences may be considered:

$$H_0(G): \bar{F}_{1\bullet} = \bar{F}_{2\bullet} =,...,= \bar{F}_{a\bullet}, \tag{3}$$

which means that the distribution F_{ij} over the t periods of time is the same for each of the a groups.

2. Let $\bar{F}_{\bullet j} := E_j(F) = (\sum_{i=1}^{a} F_{ij})/a$ denote the average values of the distribution F_{ij} for the j-th time point over all groups. We should test the $H_0(T)$-hypothesis of no relative time-associated effect T (time factor) for the groups compared

$$H_0(T): \bar{F}_{\bullet 1} = \bar{F}_{\bullet 2} =,...,= \bar{F}_{\bullet t} \ , \tag{4}$$

in other words, to test whether the underlining time trend for the observed variable is the same in the groups being compared.

3. Let $\bar{F}_{\bullet\bullet} := E(\bar{F}_{i\bullet}) = E(\bar{F}_{\bullet j}) = (\sum_{j=1}^{t}\sum_{i=1}^{a} F_{ij}) / N$. Now, let $\delta F_{ij} = F_{\bullet j} + F_{i\bullet} - F_{\bullet\bullet}$ denote the distribution function, which allows us to test the $H_0(GT)$-hypothesis of no interaction between the time-associated effect T (time factor) and the group-associated effect (group factor) for the groups compared. There is a no interaction hypothesis between both factors

$$H_0(TG): \quad \delta F_{ij} = 0 \quad ; \quad i = 1,...,a \, , \, j = 1,...,t \tag{5}$$

or, in matrix notation, this hypothesis is written as

$$H_0(GT): \quad \delta\mathbf{F} = \begin{pmatrix} F_{11} - \bar{F}_{1\bullet} - \bar{F}_{\bullet 1} + \bar{F}_{\bullet\bullet} \\ \vdots \\ F_{a1} - \bar{F}_{a\bullet} - \bar{F}_{\bullet t} + \bar{F}_{\bullet\bullet} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{0} \, .$$

4. We will test the $H_0(T \mid G)$-hypothesis of no simple time effect within the groups

$$H_0(T \mid G): F_{i1} = F_{i2} =,..., F_{it} = \bar{F}_{i\bullet} \tag{6}$$

2.2.2. Estimator of the relative treatment effect

Since no parameters are involved in the nonparametric model (1)-(2), the distribution functions $F_{ij}(x)$ are used to describe an effect (e.g. time-related or treatment group-related effect) based on the rank-value estimates. Using the rank approach we can apply the (1)-(2) for censored data (or baseline data points) treated by the model as 'tied' ranks. Let R_{ikj} denote the (mid)-rank of the random variable X_{ikj} among all the N observations. Then for the i-*th* group ($i = 1,...,r$) at the time-point j ($j = 1,...,t$) we can estimate the relative treatment effect by the formula (Akritas, Brunner, 1997)

$$\hat{q}_{ij} = n_i^{-1} \sum_{k=1}^{n_i} \frac{1}{N}(R_{ikj} - \frac{1}{2}) \quad . \tag{7}$$

Then, the vector $\hat{\mathbf{q}}_i = (\hat{q}_{i1},...,\hat{q}_{it})'$ enables a simple presentation of the time-course dynamics of the relative rank for the i-*th* treatment group. Eq.(7) provides the unbiased estimator of average rank of each group at each point in time.

2.2.3. Software

Our software was developed using Pascal and Fortran-90 (V.S.S . and V.A.K.) and is available by request.

3. MAJOR FORMULAS FOR THE NON-PARAMETRIC ANOVA ALGORITHMS

3.1. The Mack Two-Way ANOVA Test (no interactions between factors)

The Mack two-way ANOVA test is a generalization of the Friedman and the Kruskal-Wallis rank tests for a two-factor experiment in a scheme with randomized blocks (Mack, 1981). But if, in the Friedman model, there is only one observation that can lie within each i-th block for the j-th treatment, whereas in our case, with repeated measurements, more than one observation can lie in some blocks. In essence, Mack's test is a sum of the Kruskal-Wallis rank statistics that has been taken on all blocks available. In our example, the role of ‘block factor’ is assigned to a particular vaccination method (called G) applied to a monkey, and the role of ‘treatment’ is assigned to a time factor (called T). The approach by *Mack* to the 2-way ANOVA consists of the following. The initial data are the 2-D table with $N = n \cdot t$ observations classified on a groups (or treatments) of subjects, where $n = \sum_{i=1}^{a} n_i$, and n_i is a size of i-th group. The model of data is

$$X_{ij} = \mu + \beta_s + \tau_i + \varepsilon_{is}, \quad s = 1 ,..., t \;\; , \quad i = 1 ,..., a$$

where μ is a (unknown) mean value of variable X for the subject , τ_i – is an (unknown) effect of i-th treatment $\left(\sum_{i=1}^{a} \tau_i = 0 \right)$, β_s – is an (unknown) effect of s-th block $\left(\sum_{s=1}^{t} \beta_s = 0 \right)$ and ε_{is} is a random error. In the model, it is assumed that all errors ε_{is} are mutually independent and each random value ε_{is} has got the same continuous (unknown) distribution function. (In practice, the assumption of the continuity means that ties in rank values were not allowed or appeared rarely).

3.2. The Brunner- Akritas-Puri's models (non-parametric repeated measures ANOVA)

3.2.1. *Notations and definitions*

By the mixed model, a different treatment groups are considered, where every i-th group ($i=1,..., a$) contains independent and randomly chosen subjects. These $n = \sum_{i=1}^{a} n_i$ subjects are observed repeatedly under $s = 1,...,t$ different situations, or treatments (generally t time points). In our investigation, we applied so-called *split-plot* design of repeated measures (Brunner & Puri, 2001).

We described below *only estimating formulas* of the BAP tests, that we used to build-up our software. Besides, we re-written the formulas in more simple forms for programming and tested the formulas using data presented in (Brunner, Puri, 2001; Akritas, Brunner, 1997) be sure the results coincide with the test data.

Let denote $N = n \cdot t$ the total number of observation (with no missing data). Then we introduce the value

$$R_{\bullet\bullet\bullet} = \sum_{i=1}^{a} \sum_{k=1}^{n_i} \sum_{s=1}^{t} R_{iks}$$

of the mean rank, where R_{iks} are rank elements of array $\mathbf{R}$. The value R_{iks} is the mid-rank of X_{iks} among all N observations from a data matrix $\mathbf{X}$. We will write elements of the $\mathbf{R}$ with three indexes: $\mathbf{R} = \left(R_{iks} \right)$, $i = 1,...,a$; $s = 1,...,t$; $k = 1,...,n_i$; where $a \geq 2$ is the number of groups, t $(t = 2, 3,...)$ is the number of time-points, and n_i ($n_i = 2, 3,...$) is the number of subjects in i-th group.

We use the following auxiliary matrixes and vectors (or its elements):

- the matrix $\mathbf{R}_{\mathbf{ks}}^{(i)} = \left(\mathbf{R}_{\mathbf{i1}}^{T}, \ldots, \mathbf{R}_{\mathbf{in_i}}^{T} \right)^{T}$ with the vector-strings $\mathbf{R}_{\mathbf{ik}}^{\mathbf{T}} \in \Re^{\mathbf{t}}$ contained the ranks R_{iks} ;
- the matrix $\mathbf{R}_{\mathbf{ik\bullet}} = \left(R_{ik\bullet} \right)$ with the elements $R_{ik\bullet} = t^{-1} \cdot \sum_{s=1}^{t} R_{iks}$;
- the matrix $\mathbf{R}_{\mathbf{i\bullet s}} = \left(R_{i\bullet s} \right)$ with the elements $R_{i\bullet s} = n_i^{-1} \cdot \sum_{k=1}^{n_i} R_{iks}$;

- the vector $\mathbf{R}_{\mathbf{i}\bullet\bullet} = \left(R_{1\bullet\bullet}, \ldots, R_{a\bullet\bullet}\right)^T$ of the coordinate space $\Re^a$ with the elements $R_{i\bullet\bullet} = n_i^{-1} \cdot \sum_{k=1}^{n_i} R_{ik\bullet}$;
- the vector $\mathbf{R}_{\bullet\bullet s} = \left(R_{\bullet\bullet 1}, \ldots, R_{\bullet\bullet t}\right)^T \in \Re^t$ consisting of the elements
$$R_{\bullet\bullet s} = \left(an_i\right)^{-1} \cdot \sum_{i=1}^{a} \sum_{k=1}^{n_i} R_{iks} = a^{-1} \cdot \sum_{i=1}^{a} R_{i\bullet s} \quad .$$

The matrix $\mathbf{R}_{\mathbf{i}\bullet\mathbf{s}} = \left(R_{i\bullet s}\right)$ also one is used in another form, by its vector-strings $\mathbf{R}_{\mathbf{i}\bullet\mathbf{s}} = \left(\mathbf{R}_{1\bullet}^T, \ldots, \mathbf{R}_{\mathbf{a}\bullet}^T\right)$ where the vector-column $\mathbf{R}_{\mathbf{i}\bullet}$ contains the mean ranks in i-th group G (or treatment).

Now we provide the asymptotic estimates for testing of four null-hypotheses.

3.2.2. The treatment (group-associated) effect $H_0(G)$

We used the statistic
$$F_n\left(A\right) = \frac{a}{d\left(a-1\right)} \cdot \sum_{i=1}^{a} \left(R_{i\bullet\bullet} - R_{\bullet\bullet\bullet}\right)^2 , \tag{8}$$
where $d = \sum_{i=1}^{a} \left(\hat{\sigma}_i^2 / n_i\right)$, where $\hat{\sigma}_i^2 = \left(n_i - 1\right)^{-1} \cdot \sum_{k=1}^{n_i} \left(R_{ik\bullet} - R_{i\bullet\bullet}\right)^2$.

Under the null hypothesis, the distribution $F_n\left(A\right)$ has approximately the central $\mathrm{F}\left(\hat{f}_G, \infty\right)$ -distribution with an estimation of degree of freedom
$$\hat{f}_G = \left(a-1\right)^2 \cdot \left\{1 + a\left(a-2\right) \cdot d^{-2} \cdot \sum_{i=1}^{a} \left(\hat{\sigma}_i^2 / n_i\right)^2 \right\}^{-1} \tag{9}$$

It is easy to see that $\hat{f}_G = 1$ if $a = 2$. This approximation is rather accurate for the small samples and provides slightly conservative estimates. Eqs.(8)-(9) were recommended as an estimation of the true distributions for small groups of subjects having observations in a few time-points (Brunner, Munzel, Puri, 1999).

3.3.3. Time-associated effect $H_0(T)$

We used the statistic
$$F_n\left(t\right) = \left(n / g\right) \cdot \sum_{s=1}^{t} \left(R_{\bullet\bullet s} - R_{\bullet\bullet\bullet}\right)^2 , \tag{10}$$
which has asymptotically a central $\mathrm{F}\left(\hat{f}_T, \infty\right)$ distribution, where $g = trace\left(\mathbf{P_t} \cdot \hat{\mathbf{V}}_t\right)$ with the degree of freedom
$$\hat{f}_T = g^2 \Big/ trace\left[\mathbf{P_t} \cdot \hat{\mathbf{V}}_\mathbf{t}\right]^2 \tag{11}$$
and $t \times t$ -symmetric matrix $\mathbf{P_t}$ consists of elements $-1/t$ out of its major diagonal cells and with $\left(1 - 1/t\right)$ placing into the diagonal cells, and matrix
$$\hat{\mathbf{V}}_\mathbf{t} = a^{-2} \cdot \sum_{i=1}^{a} \hat{\mathbf{V}}_\mathbf{i} \quad ,$$
where

$$\hat{\mathbf{V}}_{\mathbf{i}} = \left[n \big/ n_i(n_i-1) \right] \cdot \sum_{k=1}^{n_i} \mathbf{e}_{ik}\mathbf{e}_{ik}^{T} \quad ,$$

where the vector $\mathbf{e}_{ik} = \mathbf{R}_{\mathbf{ik}} - \mathbf{R}_{\mathbf{i}\bullet\bullet}$ belong to the coordinate space $\Re^t$, and the vector $\mathbf{R}_{\mathbf{ik}}$ is *k*-th transposed string of the $n_i \times t$-rank matrix $\mathbf{R}_{\mathbf{ks}}^{(i)}$, and $\mathbf{R}_{\mathbf{i}\bullet\bullet}$ denote the mean of these rank vectors within *i*-th group *G*.

3.2.4. The interaction test $H_0(GT)$

We used the statistic

$$F_n(\mathbf{C}_{\mathbf{GT}}) = n \cdot c^{-1} \cdot \sum_{i=1}^{a}\sum_{s=1}^{t}\left(R_{i\bullet s} - R_{i\bullet\bullet} - R_{\bullet\bullet s} + R_{\bullet\bullet\bullet}\right)^2 , \tag{12}$$

which has asymptotically a central $\mathrm{F}\left(\hat{f}_{GT}, \infty\right)$ distribution, where $c = trace\left(\mathbf{C}_{GT} \cdot \hat{\mathbf{V}}_n\right)$ with the degree of freedom

$$\hat{f}_{GT} = c^2 \big/ trace\left[\mathbf{C}_{GT} \cdot \hat{\mathbf{V}}_n\right]^2 \quad , \tag{13}$$

where $\mathbf{C}_{\mathbf{GT}}$, $\hat{\mathbf{V}}_n$ are the symmetric block-type square matrixes of size $at \times at$:

$$\mathbf{C}_{\mathbf{GT}} = \mathbf{P}_a \otimes \mathbf{P}_{\mathbf{t}} \quad ,$$

where $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker-product (direct product) of some matrixes $\mathbf{A}$, $\mathbf{B}$ and matrix $\mathbf{P}_{\mathbf{t}}$ was defined above, $a \times a$-symmetric matrix $\mathbf{P}_a$ consists of by the same way as the matrix $\mathbf{P}_{\mathbf{t}}$ but with $-1/a$ and $\left(1 - 1/a\right)$ instead of the elements $-1/t$ and $\left(1 - 1/t\right)$, and matrix $\hat{\mathbf{V}}_n$ has got a block-and-diagonal structure $\hat{\mathbf{V}}_{\mathbf{n}} = diag\left(\hat{\mathbf{V}}_{\mathbf{1}}, \ldots, \hat{\mathbf{V}}_{\mathbf{a}}\right)$ with its non-zero block-and-diagonal matrix $\hat{\mathbf{V}}_i$ $(i = 1, \ldots, a)$ of size $t \times t$ that has been already defined above (see subsection *Time-associated effect $H_0(T)$*).

Thus, a block-type element of matrix $\mathbf{C}_{\mathbf{GT}}$ is a $t \times t$-matrix $\left(-1/a\right) \cdot \mathbf{P}_{\mathbf{t}}$ out of its block-and-diagonal elements and all the last comprises of $t \times t$-matrix $\left(1 - 1/a\right) \cdot \mathbf{P}_{\mathbf{t}}$.

Starting from the matrix form (Brunner, Puri, 2001), we derived an explicit formula

$$c = trace\left(\mathbf{C}_{GT} \cdot \hat{\mathbf{V}}_n\right) = \left(1 - 1/a\right) \cdot \left[trace\left(\hat{\mathbf{V}}_{\mathbf{n}}\right) - t^{-1} \sum_{i=1}^{a}\sum_{j=1}^{t}\sum_{s=1}^{t} v_{s,j,i} \right], \tag{14}$$

where $v_{s,j,i}$ is the element of matrix $\hat{\mathbf{V}}_i$ on the crossing of its *s*-th row and *j*-th column and $v_{s,s,i}$ is its *s*-th diagonal element. We also have derived the relatively simple explicit expression for $trace\left[\mathbf{C}_{GT} \cdot \hat{\mathbf{V}}_n\right]^2$, however for our needs the $trace\left[\mathbf{C}_{GT} \cdot \hat{\mathbf{V}}_n\right]^2$ was calculated by straight way via the aforesaid $at \times at$-matrix.

4. The conditional test $H_0(\mathrm{T|G})$

We used the χ^2-type statistic

$$Q_n(T/A) = \sum_{i=1}^{a}\left[\sum_{j=1}^{t}\sum_{l=1}^{t} R_{i\bullet j} \cdot \hat{s}_{jl}^{(i)} \cdot R_{i\bullet l} - \frac{1}{\hat{s}_{\bullet\bullet}^{(i)}}\left(\sum_{j=1}^{t} R_{i\bullet j} \cdot \hat{s}_{j\bullet}^{(i)} \right)^2 \right] , \tag{15}$$

(Akritas, Brunner, 1997) and $\hat{s}_{jl}^{(i)}$ is (j,l)-th element of inverse matrix $\tilde{\mathbf{V}}_{\mathbf{i}}^{-\mathbf{1}}$, where $\tilde{\mathbf{V}}_{\mathbf{i}} = \hat{\mathbf{V}}_{\mathbf{i}} / t$. The matrix $\hat{\mathbf{V}}_{\mathbf{i}}$ is the block in a block-and-diagonal structure $\hat{\mathbf{V}}_{\mathbf{n}} = diag\left(\hat{\mathbf{V}}_{\mathbf{1}}, \ldots, \hat{\mathbf{V}}_{\mathbf{a}}\right)$ that has been already defined above (see subsection *Time-associated effect $H_0(T)$*). We also used notations:

$$\hat{s}_{\bullet\bullet}^{(i)} = \sum_{j=1}^{t}\sum_{l=1}^{t}\hat{s}_{jl}^{(i)}, \quad \hat{s}_{j\bullet}^{(i)} = \sum_{l=1}^{t}\hat{s}_{jl}^{(i)} \quad .$$

The statistic (15) has asymptotically the χ^2 distribution with the $f = a \cdot (t-1)$ degrees of freedom.

3.3.5. Relative marginal effect p_{is} and its asymptotically unbiased estimator $\hat{p}_{is}$

The marginal distribution functions $F_{is}(x)$ are used to describe the relative treatment effect. The *relative marginal (treatment) effect* function is

$$p_{is} = \int H(x) \cdot dF_{is}(x),$$

where $H(x) = N^{-1} \cdot \sum_{i=1}^{a}\sum_{s=1}^{t} n_i F_{is}(x)$ is the average of all N normalized distribution functions observed in our data.

The relative marginal effects p_{is} are estimated by replacing $F_{is}(x)$ and $H(x)$ by their empirical functions (see details in Brunner, Puri, 2001). As a result, the estimator $\hat{p}_{is}$ will be equaled to

$$\hat{p}_{is} = n_i^{-1} \cdot \sum_{k=1}^{n_i} \frac{1}{N}\left(R_{iks} - 1/2\right) \quad .$$

(see Eq.(7) in Methods section). It was shown that the estimator $\hat{p}_{is}$ is $\mathbf{L}_{\mathbf{2}}$-consistent for p_{is}, i.e. $\mathrm{E}(\hat{\mathrm{p}}_{\mathrm{is}} - p_{is})^2 \to 0$ as $\min\{ n_i \} \to \infty$, $i = 1,\ldots,a;\; s = 1,\ldots,t$.

4. RESULTS

4.1. Estimating the Relative Treatment Effects of Vaccinations on Virus Load

Figures 1a,b,c show the dynamics of the virus load in plasma of monkeys in groups A (immunized IR with adjuvant alone), B (immunized SC with HIV/SIV peptide vaccine) and C (immunized IR with HIV/SIV peptide vaccine).

Goodness-of-fit statistical analysis showed that the concentration of SHIV mRNA in the circulating blood dropped in all groups with similar half-lives ($t_{1/2}$ of 10 days, 9.8 days and 9.2 days in groups C, B and A, respectively) between 21 and 63 days; however, in macaques immunized intrarectally (group C), viral mRNA titer was reduced to an undetectable level by ~ 63rd - day, and remained undetectable subsequently (Figure 1c). At day 63 after challenge, early in the set-point (the quasi steady state phase of infection process), groups B and C were statistically different according to the Wilcoxon rank-sum test ($p < 0.05$). The viral loads at all time points measured between days 49 to 196 showed a significant difference between macaques immunized intrarectally (group C) and immunized subcutaneously (group B) by the BAP test (see also Table 2a,b and Table 4).

The BAP model shows no a synergy between group-associated and time-associated effects (Table 2c). This last result allows us to compare the significance values provided by the BAP test and by Mack's test (which ignores interactions between the two factors). Tables 3, 4 and 5 show that there is a qualitative agreement between the estimates made by these two models.

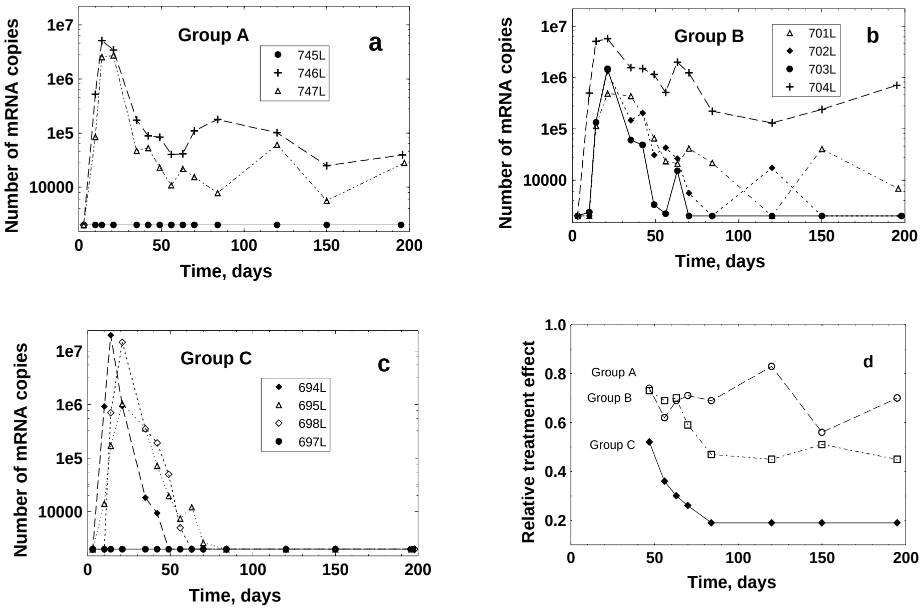

Figure 1. Dynamics of virus load in groups A, B and C (panels a, b, c) and time curves of the estimated relative treatment effects for the three groups (the two animals without viremia were excluded) (panel d)

Table 2. Virus load. The Brunner-Akritas-Puri (BAP) ANOVA tests. t= 8 time points between 49 and 196 days after SHIV challenge. Bold numbers: $p \leq 0.05$

a: the test of no group-related (*G*) effect and the test of no time-related (*T*) treatment effect; *n*=9 (total number of monkeys with viremia)

Table 2a

Effect	Groups	*n*	*F*	$\hat{f}$	*P*
G	A vs B vs C	2+4+3	4.48	1.62	**0.017**
T			3.58	3.22	**0.011**
G	B vs C	4 +3	4.67	1.0	**0.031**
T			4.75	2.82	**0.003**
G	A vs C	2+3	15.79	1.0	**<0.001**
T			2.19	2.78	0.092
G	A vs B	2+4	0.2	1.0	0.654
T			1.36	1.90	0.256

b: the test of no group-related (G) effect and the test of no time-related (*T*) treatment effect; *n*=11

Table 2b

Effect	Groups	*n*	*F*	$\hat{f}$	*P*
G	A vs B vs C	3+4+4	1.77	1.58	0.178
T			3.74	2.99	**0.011**
G	B vs C	4 +4	6.07	1.0	**0.014**
T			4.38	2.60	**0.007**
G	A vs C	3+4	2.17	1.0	0.141
T			1.95	2.52	0.131
G	A vs B	3+4	0.06	1.0	0.800
T			1.62	2.38	0.193

c: test of no interaction between the group-related and time-related treatment effects (*GT*); n=9 (total number of monkeys with viremia).

Table 2c

Effect	Groups	*n*	F	$\hat{f}$	*P*
GT	A vs B vs C	2+4+3	1.498	4.022	0.194
GT	B vs C	4 +3	0.276	2.815	0.830
GT	A vs C	2+3	2.036	2.783	0.111
GT	A vs B	2+4	2.231	1.901	0.111

Table 3. Virus load. The Mack two-way ANOVA test (see details in Table 2 above).

a: tests of no main treatment effect; *n*=9 (total number of monkeys with viremia)

Table 3a

Groups	*n*	F	f_1 / f_2	*P*
A vs B vs C	2+4+3	2.401	16/48	**0.001**
B vs C	4 +3	2.707	8/ 40	**0.018**
A vs C	2+3	17.21	8/ 32	**<0.001**
A vs B	2+4	0.28	8/ 32	> 0.9

b: test of no main treatment effect; *n*=11

Table 3b

Groups	*n*	F	f_1 / f_2	*p*
A vs B vs C	3+4+4	2.057	16/64	**0.022**
B vs C	4 +4	4.105	8/ 48	**0.001**
A vs C	3+4	2.979	8/ 40	**0.010**
A vs B	3+4	0.239	8/ 40	> 0.9

Table 4. Robustness of significances of the tests by BAP (H_0 *(T)* and H_0 *(G)*) and by the Mack test ($H_0(G)$) relatively the number of repeated measures in time used for our analysis. The robustness was evaluated by a significance of the p-values (bold numbers; $p \leq 0.05$) of the two-side test, when data for consequence time points starting from day 195 and up to day 56 (after SHIVKu2 virus challenge) were progressively excluded from analysis.

a: three groups: *n*= 2+4+3=9 in the A vs B vs C comparison

Table 4a

Since Day /Times	F, BAP *Ho(T)*	$\hat{f}$	*p*	F, BAP *Ho(G)*	$\hat{f}$	*p*	F, Mack	f_1 / f_2	*p*
195/9	6.433	3.218	**<0.001**	3.490	1.677	**0.0384**	2.106	18/54	**0.0183**
150/8	6.358	2.892	**<0.001**	3.028	1.736	0.0558	2.017	16/48	**0.0313**
120/7	5.142	2.502	**0.0029**	2.881	1.773	0.0627	2.050	14/42	**0.0369**
84/6	6.275	2.124	**0.0015**	2.134	1.799	0.1239	1.927	12/36	0.0640
70/5	6.314	2.063	**0.0016**	1.897	1.864	0.1531	1.881	10/30	0.0885
63/4	8.052	1.288	**0.0021**	1.506	1.925	0.2225	1.740	8/24	0.1402
56/3	15.50	1.084	**<0.001**	0.866	1.947	0.4179	1.045	6/18	0.4297

b: two groups: n= 4+3=7 in the B vs C comparison

Table 4b

Since Day /Times	F, BAP *Ho(T)*	$\hat{f}$	p	F, BAP *Ho(G)*	$\hat{f}$	p	F, Mack	f_1/f_2	P
195/9	7.668	2.792	<**0.001**	4.053	1	**0.0441**	2.497	9/45	**0.0208**
150/8	7.343	2.781	<**0.001**	4.084	1	**0.0433**	2.556	8/40	**0.0237**
120/7	7.593	2.70	<**0.001**	3.928	1	**0.0475**	2.632	7/35	**0.0270**
84/6	7.226	2.010	<**0.001**	3.657	1	0.0558	2.737	6/30	**0.0305**
70/5	5.522	1.626	**0.0071**	3.365	1	0.0666	2.889	5/25	**0.0342**
63/4	4.546	1.159	**0.0276**	2.813	1	0.0935	2.911	4/20	**0.0476**
56/3	9.841	1.117	**0.0011**	1.560	1	0.2118	1.545	3/15	0.2439

c: two groups: n= 2+3=5 in the A vs C comparison

Table 4c

Since Day /Times	F, BAP *Ho(T)*	$\hat{f}$	p	F, BAP *Ho(G)*	$\hat{f}$	p	F, Mack	f_1/f_2	P
195/9	4.384	2.638	**0.0063**	10.554	1	**0.0012**	7.781	9/27	<**0.001**
150/8	4.612	2.515	**0.0055**	8.517	1	**0.0035**	6.828	8/24	**0.0001**
120/7	3.058	2.372	**0.0383**	7.342	1	**0.0067**	5.826	7/21	**0.0007**
84/6	3.981	1.818	**0.0220**	4.629	1	**0.0314**	4.769	6/18	**0.0045**
70/5	5.035	1.586	**0.0113**	3.606	1	0.0576	3.654	5/15	**0.0232**
63/4	6.459	1.112	**0.0089**	2.358	1	0.1247	2.871	4/12	0.0700
56/3	10.338	1.084	<**0.001**	1.318	1	0.2510	1.909	3/9	0.1986

The 'group-related treatment effect' is the main effect we are testing. Figure 1d (see above) shows the time-course dynamics of the estimated relative treatment effect for the studied groups during the quasi-steady state (chronic) phase of the infection process for the 9 out of 11 monkeys that actually demonstrated viremia, measured using standard NASBA assay for viral RNA. Figure 1d also shows that the estimated relative treatment effect function in group C has occurred (kinetically) faster and that the viral load falls much lower than in group B.

Table 5. Robustness of significances of the tests by BAP (*Ho(T)* and *Ho(G)*) and by the Mack test (*Ho(G)*) relatively the number of repeated measures in time used for our analysis. The robustness was evaluated by a significance of the p-values (bold numbers; $p \leq 0.05$) of the two-side test, when data for consequence time points starting from day 42 day up to day 120 after SHIVKu2 virus challenge were excluded from analysis.

a: three groups: *n*= 2+4+3=9 in the A vs B vs C comparison

Table 5a

Since Day /Times	F, BAP *Ho(T)*	$\hat{f}$	p	F, BAP *Ho(G)*	$\hat{f}$	p	F, Mack	f_1/f_2	p
42/10	10.507	2.983	<**0.001**	2.951	1.705	**0.061**	1.917	20/60	**0.028**
47*/9	6.433	3.218	<**0.001**	3.490	1.677	**0.038**	2.106	18/54	**0.018**
49/8	3.582	3.221	**0.011**	4.481	1.618	**0.017**	2.401	16/48	**0.010**
56/7	2.250	2.601	0.090	5.696	1.528	**0.007**	2.748	14/42	**0.006**
63/6	1.888	2.279	0.145	6.097	1.439	**0.006**	2.876	12/36	**0.007**
70/5	1.057	2.093	0.350	5.988	1.297	**0.009**	2.441	10/30	**0.029**
84/4	0.316	1.507	0.666	6.651	1.220	**0.006**	2.418	8/24	**0.045**
120/3	0.785	1.111	0.388	7.388	1.066	**0.006**	2.505	6/18	**0.061**

b: two groups: n= 4+3=7 in the B vs C comparison

Table 5b

Since Day /Times	F, BAP *Ho(T)*	$\hat{f}$	p	F, BAP *Ho(G)*	$\hat{f}$	p	F, Mack	f_1 / f_2	P
42/10	11.576	2.677	< **0.001**	3.692	1	**0.055**	2.337	10/50	**0.024**
47*/9	7.668	2.792	< **0.001**	4.053	1	**0.044**	2.498	9/45	**0.021**
49/8	4.749	2.815	**0.0032**	4.667	1	**0.031**	2.707	8/40	**0.018**
56/7	3.702	2.452	**0.0172**	5.328	1	**0.021**	2.996	7/35	**0.014**
63/6	2.534	2.149	0.0752	5.061	1	**0.025**	3.085	6/30	**0.018**
70/5	1.263	1.833	0.2813	3.867	1	**0.049**	2.196	5/25	0.087
84/4	0.028	1.037	> 0.8	3.880	1	**0.049**	2.059	4/20	0.125
120/3	none	none	> 0.8	4.200	1	**0.040**	2.059	3/15	0.149

* virus load value was estimated by the quadratic interpolation of the virus load measurements observed on the days 42, 49 and 53.

The two "outliers" 697L, 745L of the 11 monkeys were non-infected (perhaps because of technical reasons) related to the challenge, not to any immune response. They were classified as non-infected by the act that the viral load was negative from the beginning of our observations, even the ones that were never immunized (measured by NASBA, data not shown). The blood cell counts (including for CD4(+) cells and CD8(+) cells) for the monkeys 697L and 745L did not have also a typical dynamics of immune response on SHIV infection.

With regard to the primary differences in viremia ('group-related treatment effects'), groups B (subcutaneous immunization) and C (intrarectal immunization) are shown to be statistically different for the group-related treatment effect, whether one uses the set of 9 animals (Table 2a) or the set of all 11 animals (Table 2b). This is the key comparison of virus load observations that supports the major conclusion that intrarectal immunization is more effective than subcutaneous immunization for the long-term control of viremia (Figure 2).

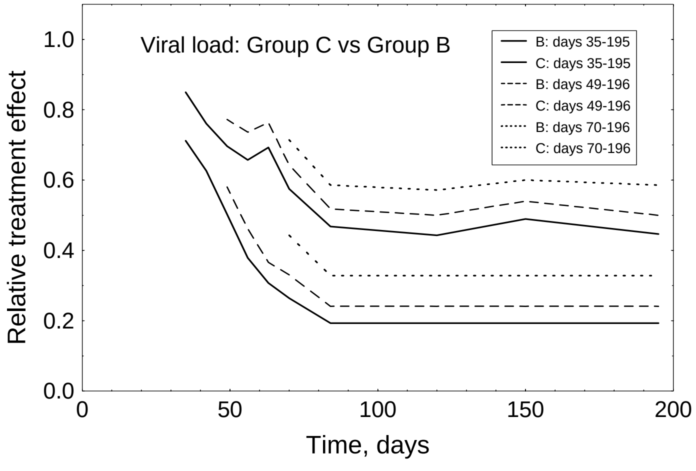

Figure 2. Virus load: Analysis of reliability of the estimates of the differences between groups B and C (see text for details)

The comparisons of groups A, B and C are shown to be statistically different from each other. The pair-wise comparisons reveal that group C is statistically different from the group A by the group treatment effect when only the viremic animals are considered, and shows the opposite trends for these groups in time when the two animals that did not become obviously infected were included (Table 2a,b). The group C vs the group A comparison is affected most by the two animals that did not show viremia, because one such animal occurs in each of these groups. However, since two macaques were likely not be infected, we think that the more relevant comparison is among the group of 9 macaques; the analysis of all 11 animals is included for completeness. In any case, the key comparison is between the groups B and C, which we show to be significant. The group B (subcutaneous immunization) is not significantly different from the control group A (adjuvant only) in either case, again consistent with the conclusion that subcutaneous immunizations with HIV/SIV peptide vaccine produced a less protective effect against pathogenic SHIV-Ku2 than did the intrarectal immunization with the same vaccine construct.

The “time-related effects” are also significant when comparing groups C vs B (p=0.003 in Table 2a, p=0.007 in Table 2b). In particular, the slopes indicating the kinetics of the viral decay curve for the considered time interval are significantly different for the compared groups.

Note that there were no missing values on our data for all observed periods of time. However, our analysis showed high robustness to missing values. The differences between virus loads for groups A, B and C remained statistically significant (at the level $p \leq 0.05$) even when up to 65% of the possible time-points have been artificially excluded from the comparative analysis (Table 4 and 5). For example, starting from days 42, 49, 56, 63, 70, 84 and 120, we used virus load data sets from 9, 8, 7, 6, 5, 4 and 3 time-points, respectively (Table 5a). Using the BAP test for $H_0(G)$, we found that starting from day 47, all of the shorter time intervals indicate that there are significant differences between groups A, B and C (Table 5a). The same result was observed for the pair-wise comparison of groups B and C (Table 5b). The Mack test also showed a quite robust p-value when the number of time points available for analysis were reduced (Tables 4 and 5). Figure 2 shows time curves of the estimated relative treatment effects for the two vaccinated groups B and C when different numbers of time-points, starting from day 35, 49 and 70, were used for our comparative analyses. This figure shows that the main treatment effect is a quite robust to the number of time points taken for our analyses and virus load data on a few time-points starting from day 47 (or so around day 49) demonstrates significant differences between groups A, B and C. .

Hence, even though the precision of the estimated p-values is limited by the small sample size, we believe that the conclusions presented in this section are accurate and robust.

4.2. Relative Treatment Effects of the Vaccinations on the CD4(+) and CD8(+) Cells in the Blood

Figures 3 and 4 show the dynamics of the absolute numbers of CD4(+) and CD8(+) T cells in all peripheral blood specimens taken at various times after viral challenge in the macaques in groups A, B and C. The following question of interest to an immunological assessment of vaccine efficacy is addressed: is there a significant relative improvement in the dynamics of the CD4(+) and CD8(+) cells at the chronic infectious phase that is associated with the local (IR) method versus global (SC) methods of vaccination?

We compared the relative treatment effects on CD4(+) cell, starting on day 35 and ending on day 176 (t=5 time points). Data from the 9 monkeys with viremia have been analyzed. The relative counts of CD4(+) cells in monkeys of group C are mostly higher than those of group A or of group B (Figure 3d). Statistically significant differences we observed by the group-associated effects in the comparisons of the groups A vs B vs C, which reflects mostly the differences between groups A and C (Table 6a).

Significant differences (p <0.011) were also observed in the time-associated effects in these comparisons of the groups: A vs B vs C; B vs C; A vs C, but no differences were observed for the comparisons between groups A and B (Table 6a). It is interesting that it is these two groups that show a significant interaction effect (p<0.001), which reflects a significant trend in the number of CD4(+) cells in blood of monkeys of the IR vaccinated group C. Similar statistical differences we found when data from all of the 11 monkeys were included in the analysis (not presented). In agreement with the BAP test (without interaction effect), the Mack’s test shows a significant and less conservative estimate of significance level (Table 6b).

In the case of CD8(+) cells, the results of our comparisons of the relative treatment effects in groups A, B and C are presented in Table 7.

As described above for CD4(+) cells, we studied the CD8(+) cell counts on days 35, 99, 120, 149 and 176 (t=5 time points) in the 9 monkeys with viremia. Figure 4d (see above) shows that the estimates of the relative treatment effect on CD8 cells in monkeys of group C are higher than those of group A or of group B. Testing the null-hypotheses $H_0(G), H_0(T)$ and $H_0(GT)$, we found significant differences only by

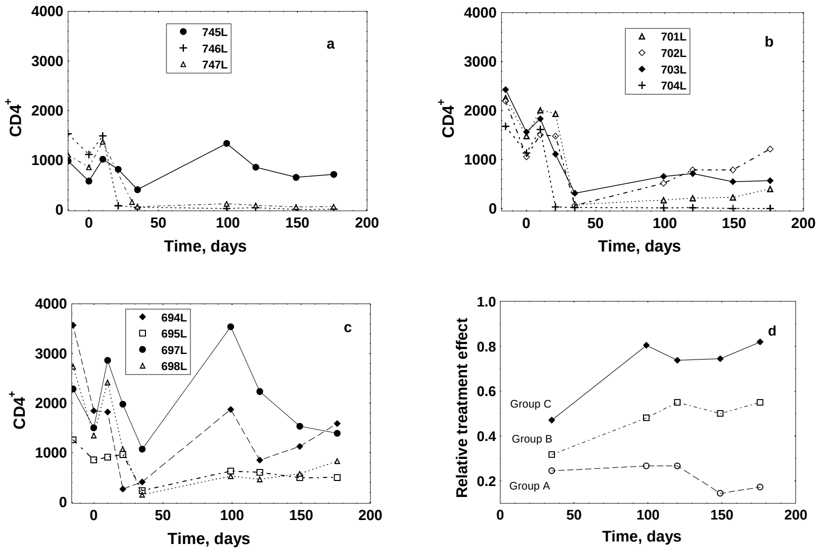

Figure 3. Dynamics of the CD4(+) cells.
a: IR immunization of macaques with mucosal adjuvant LT (R192G) alone;
b: Subcutaneous immunization of macaques with HIV/SIV peptide vaccine in Montanide ISA 51 adjuvant;
c: Intrarectal immunization of macaques with HIV/SIV peptide vaccine mixed with LT (R192G);
d: time curves of the estimated relative treatment effects of the three groups (the two animals without viremia were excluded)

Table 6. Comparison of CD4(+) cells.

a: The BAP tests of no group-related treatment effect and of no time-related effect. Data of five time points sense 35 day to 176 day after SHIV-Ku2 have been used.

Table 6a

Effect	Groups	N	F	$\hat{f}$	P
G T	A vs B vs C	2+4+3	5.5213 5.9144	1.6189 1.7728	**0.0072** **0.004**
G T	B vs C	4 +3	1.9410 7.0543	1.0 1.7814	0.1636 **0.0014**
G T	A vs C	2+3	17.089 4.5392	1.0 1.9797	**< 0.0001** **0.0110**
G T	A vs B	2+4	1.6792 1.9177	1.0 1.3276	0.195 0.1616

b: The Mack two-way ANOVA test. The same 5 time points sense 35 day to 176 day after SHIV-Ku2 challenge

Table 6b

Groups	n	F	f_1 / f_2	p
A vs B vs C	2+4+3	2.294	10/30	**0.0386**
B vs C	4 +3	1.452	5/ 25	0.241
A vs C	2+3	9.0	5/ 15	**<0.001**
A vs B	2+4	0.828	5/ 20	0.545

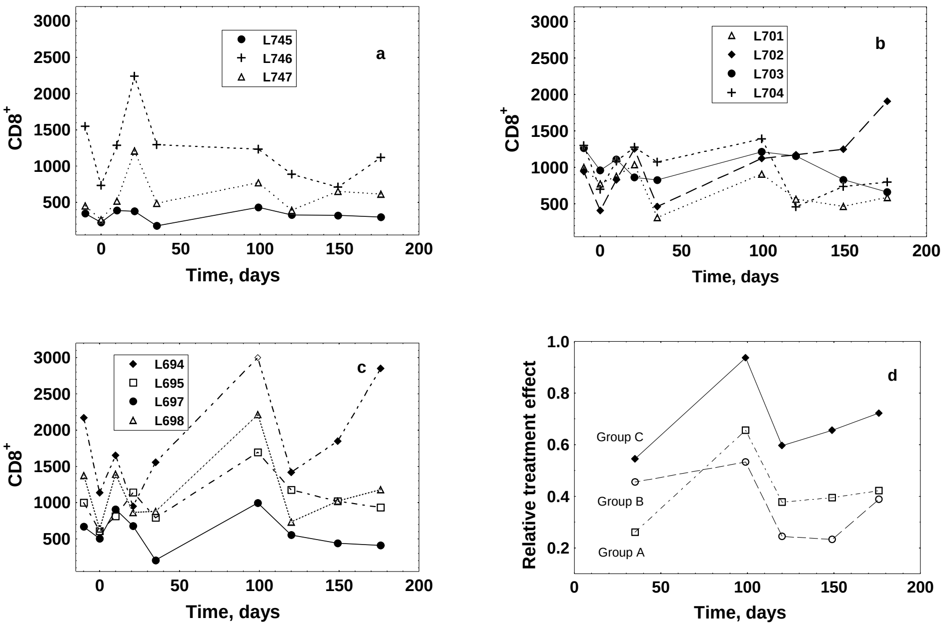

Figure 4. Dynamics of CD8(+) cells.
a: IR immunization with mucosal adjuvant LT (R192G) alone;
b: Subcutaneous immunization of macaques with HIV/SIV vaccine mixed with LT(R192G);
c: Intrarectal immunization of macaques with HIV/SIV vaccine mixed with LT(R192G);
d: time curves of the estimated relative treatment effects of the three groups (the two animals without viremia were excluded)

Table 7. Comparison of CD8 (+) cells.

a: The Brunner-Akritas-Puri (BAP) ANOVA tests. Tests of no group-related treatment effect and of no time-related effect. Five time points between 35 and 176 days have been used. Data of five time points sense 35 day to 176 day have been used.

Table 7a

Effect	Groups	*n*	F	$\hat{f}$	*p*
G T	A vs B vs C	2+4+3	1.5306 4.3357	1.6115 2.4673	0.210 0.0081
G T	B vs C	4 +3	3.4571 4.2578	1.0 2.3646	0.063 0.0098
G T	A vs C	2+3	2.2006 3.8293	1.0 1.9308	0.138 0.0231
G T	A vs B	2+4	0.0370 2.1635	1.0 1.9515	> 0.8 0.1163

b: The Mack two-way ANOVA test Table 7b

Groups	*n*	F	f_1 / f_2	*p*
A vs B vs C	2+4+3	2.061	10/30	0.0616
B vs C	4 +3	3.163	5/ 25	0.0239
A vs C	2+3	2.455	5/ 15	0.0815
A vs B	2+4	0.502	5/ 20	0.77

the time-associated effects in comparisons of the groups A vs B vs C; B vs C and A vs C, which reflects no differences between groups A and B, and a significant relative vaccination effect in group C (Table 7a). The significance of the difference between group B and C was slightly smaller (p=0.063) than the generally applied default p-value (0.05). The Mack test implies that there are significant differences (p=0.024) only for a comparison of groups C and B (Table 7b).

Because the analysis of the $H_0(T \mid G)$ justifies our making the estimates the main and interaction effects tested together (Akritas, Brunner, 1997), we compared CD8(+) cells counts in the groups B and C using the given statistic for the null-hypothesis $H_0(T \mid G)$ (Eq.8A, Appendix). We found a highly significant difference (p<0.001) between groups B and C when data points for days 35,99, 120, 149 and 176 were used. Significant p-values were also observed when data for days 99, 120, 149, and 176 (p=0.0011), or data for days 120, 149, and 176 (p=0.028) was tested.

4.3. Relationships of the Virus Load, CD4(+) and CD8(+) Responses in the Courses of a Chronic Phase of SHIV Infarction after Micosal and Subcutaneous Vaccination

Figure 5 shows the time courses of the estimated group-relative treatment effects of virus load, CD4(+) cells for the SC immunized (group B) and IR immunized (group C) groups.

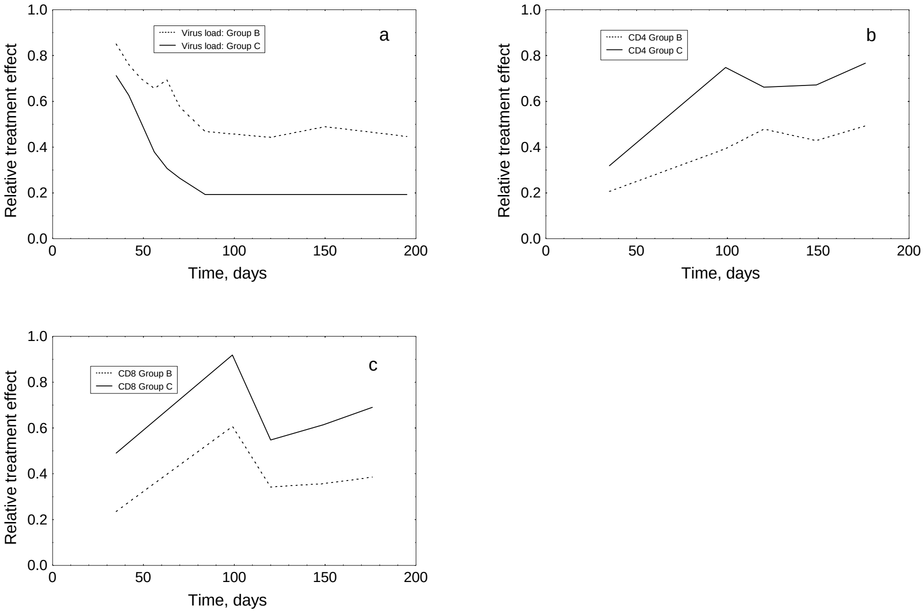

Figure 5. Comparison of the relative treatment effects between systemic immunization (Group B) and local (Group C) immunizations: **a**: virus load, **b**: the number of the CD4(+), **c**: the number of CD8 (+) cells(c).

This graphical analysis easily demonstrates a prevalence of the positive protective outcome when the method was local immunization vs systemic immunization, shows that kinetics of the CD4(+) and CD8(+) cells is quite similar, and demonstrates that these have opposite trends when compared to the kinetics of the virus load. Figures 4d and Figure 5c allow as to suggest the non-monotonous kinetics of immune response of CD8(+) cells (with picks around day 100) after SHIV exposure in vaccinated monkeys. The reciprocal relationships between the virus load and CD4(+) and CD8(+) lymphocytes agree with the negative correlations between the number of virus particles in plasma and in colon lymph nodes, on the one hand,

and the proliferative response of peripheral blood lymphocytes against PHA, on the other (Figure 6). These results suggest that exposing the subject to the peptide vaccine intrarectally significantly reduces the suppression of the proliferative responses of both the CD4(+) and CD8(+) cells in the SHIV-infected monkeys.

5. DISCUSSION

The application of modern non-parametric statistical methods can substantially improve our ability to detect biologically and clinically meaningful differences when studies include small groups with varying treatments and provide useful information about how the groups differ (Brunner, Puri, 2001; Belyakov et al., 2001; Belyakov et al., 2003). In this work, we developed the comparative analysis of small treated groups in time-dependent design experiments where observations for each subject were taken at relatively few fixed time points. We considered several non-parametric repeated measures ANOVA models (Mack, 1981; Akritas, Brunner, 1997; Brunner, Puri, 2001) to evaluate the relative efficacy of mucosal (intrarectal) and subcutaneous HIV peptide vaccines in rhesus macaques intrarectally challenged with pathogenic SHIVKu2.

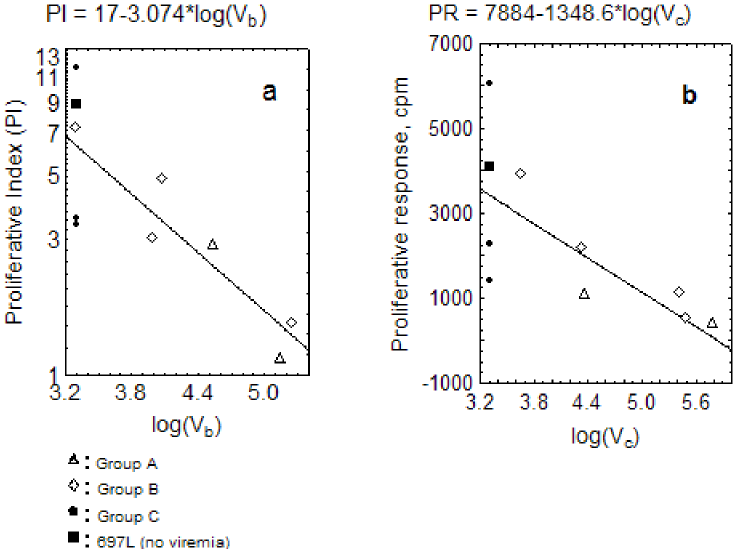

Figure 6. Proliferative response on PHA of the blood lymphocytes taken on day 95 after SHIV challenge negatively correlates with (**a**) the number of virus RNAs in plasma on the same day, and (**b**) with the number of virus RNAs in colon lymph tissues obtained at time of necropsy. Index of stimulation (IS) is the ratio of the level of proliferative response of blood lymphocytes on PHA (cpm) to spontaneous proliferative response of the lymphocytes (cpm).

Panel **a**: axis X shows log of the number of virus RNAs in the blood (V_b) of monkeys on days 90. These values were estimated by a linear interpolation between data-point on day 84 and on day 120; axis Y shows the average values of three replicated measurements of IS on day 90. Notice that virus RNA has not been detected in both the blood and lymph tissues of monkey 697L (group C). The Spearman correlation coefficient for V_b and IS is r=-0.81 ($p<0.05$).

Panel **b**: axis X shows logarithms of the number of RNAs in colon lymph tissue (V_c) obtained at time of necropsy from 10 of the 11 monkeys; no data for monkey 745L (without viremia in blood) of group A. Axis Y shows counts of the proliferative response on blood lymphocytes on PHA (cpm). Spearman correlation coefficient for r=0.85 ($p<0.05$). Solid line represents the regression function. Notice that the value of the Spearman correlation coefficient for variables V_b and V_c was also very high 0.92 ($p<0.05$).

We demonstrated with statistical significance that MamuA*01 positive macaques immunized intrarectally with synthetic HIV/SIV peptide vaccine showed a significant decline of viral titer in plasma to undetectable levels compared with subcutaneous immunized macaques and the adjuvant-only control. We also found that the numbers of the CD4(+) and CD8(+) cells of IR immunized macaques were better preserved after IR challenge with pathogenic SHIV-Ku2 than for the other two groups, and we showed that there exist strong positive trends of CD4(+) and CD8(+) cell numbers relative to the pre-steady state and steady-state phases of SHIV-Ku2 infection.

In particular, our comparisons of the estimates of the F-values (or p-values) for the different types of the relative therapeutic effects estimated by the BAP test, allows us to obtain more specific information about the mechanisms of protection provided by different vaccination methods and to compare quantitatively the impacts of the group- and time-related factors. Remarkably, in the case of CD4(+) cells, which are the major targets and reservoir for SHIV, our analysis revealed a strong synergy between these two factors.

Positive temporal dynamics of CD8(+) cells in the blood of IR-immunized monkeys could be also an important favorable prognostic factor of efficacy of immunization against SHIV infection in non-human primates. In particular, group C had a longer lag time to a decrease in viral load than did the others. On day 49 started the longest time interval after mucosal infection while IR immunization shows the statistically significant effect in viral load compared to SC immunization (by both the Mack and BAP tests, Table 5). For both CD4 and CD8 cell counts this time interval was started early on day 35 (Table 6a; Table 7a).

It is important to note that the results of our statistical tests were supported by independent clinical observation. Clinically, in each of groups A (adjuvant alone) and B (immunized with HIV/SIV vaccine subcutaneously) the animal with the highest viral load developed evidence of opportunistic infection (746L, gingivitis and pneumonia; 704L, severe gingivitis, blepharitis, facial dermatitis and pathologic evidence of pulmonary pneumocystis). None of the animals in group C (intrarectally immunized with HIV/SIV vaccine) showed evidence of these AIDS-related illnesses or other observed ailments (Belyakov et al., 2001).

Even with the small sample sizes in our study, our comparative analysis of the time course of the effect of therapeutic factors on virus titers and CD4(+) and CD8(+) cell counts in the blood of Rhesus macaques suggests that mucosal immunization of the macaques with synthetic HIV/SIV peptide vaccine was more effective than the same vaccine given subcutaneously (Belyakov et al, 2001).

Despite our results showing a relatively high significant differences between small treatment groups, we would like to stress that all BAP tests are based on some approximations of rank statistics which become better with an increasing sample size (including the number of subjects in the groups and number of different time points at which data have been observed). Since our sample was small, we compared the results obtained by BAP tests with the results obtained by Mack's ANOVA tests. If there is no interaction between the two factors, Mack's test can provide an accurate, exact estimate when the sample sizes are small (Mack, 1981). Interestingly, in our analysis the BAP test (for $Ho(G)$ hypothesis) is more conservative than the Mack test (see Tables 4,5).

We also controlled for the reliability of the estimates by demonstrating the effect of using reduced samples (Tables 4, 5). These approaches allow us to get more accurate and reliable decisions. A lot of simulation experiments have showed that the approximation formulae of the non-parametric repeated measures ANOVA-type statistics demonstrate slightly conservative estimates even relatively small sample size from the mixtures of known distribution functions have been investigated (the number of subject in the groups >6, number of groups =2,3,4 and t >3 time points; Brunner, Munzel, Puri, 1999; Brunner, Puri, 2001). However, for small sample sizes the BAP statistics should be used with caution and should be corrected in order to get a good approximation.

In this work, using Mack's and BAP models, we develop the algorithms for statistical evaluation of the relative efficacy of mucosal and subcutaneous HIV/SIV vaccines in rhesus macaques mucosally exposed to pathogenic SHIV-Ku2. We show that during the chronic phase of mucosal SHIV-Ku2 infection in immunized macaques, the numbers of the CD4(+) and CD8(+) cells reciprocally reflect the virus titers in plasma and that both parameters demonstrate that there is better protection against pathogenic SHIV-Ku2 in intrarectally immunized MamuA*01 macaques. More detail analysis of our applications of the non-parametric ANOVA methods for assessment of efficacy of the HIV/SIV peptide vaccines in rhesus macaques exposed to pathogenic SHIV-Ku2 presents in (Kuznetsov et al., 2004).

5. CONCLUSION

The general framework presented here and our analytical methodology is applicable in comparative estimates of different treatment-associated effects and their synergy for variety longitudinal data sets in studies involving small treatment groups, demonstrated here for clinical and pre-clinical studies. We expect that our approach would be valuable in many other applications that face similar limitations.

6. ACKNOLEDGMENTS

Thanks anonymous referee for providing helpful comments and recommendations.

7. REFERENCES

Akritas, M. G., and Brunner, E. (1997). A unified approach to rank tests for mixed models, *J Stat Planning and Inference* ***61***, 249-277.

Belyakov, I. M., Earl, P., Dzutsev, A., Kuznetsov, V. A., Lemon, M., Wyatt, L. S., Snyder, J. T., Ahlers, J. D., Franchini, G., Moss, B., and Berzofsky, J. A. (2003). Shared modes of protection against poxvirus infection by attenuated and conventional smallpox vaccine viruses, *Proc Natl Acad Sci U S A* ***100***, 9458-9463.

Belyakov, I. M., Hel, Z., Kelsall, B., Kuznetsov, V. A., Ahlers, J. D., Nacsa, J., Watkins, D., Allen, T. M., Sette, A., Altman, J., *et al.* (2001). Mucosal AIDS vaccine reduces disease and viral load in gut reservoir and blood after mucosal infection of macaques, *Nature Medicine* ***7***, 1320-1326.

Brunner, E., Munzel, U., Puri, M (1999). Rank-score tests in factorial designs with repeated measures, J. *Multivariate Analysis* ***70***, 286-317.

Brunner, E., and Puri, M. L. (2001). Nonparametric methods in factorial designs, *Statistical Papers* ***42***, 1-52.

Kuznetsov, V.A., Stepanov, V.S., Berzofsky, J.A., Belyakov I.M. (2004). Assessment of the relative therapeutic effects of vaccines on virus load and immune responses in small groups at several time points: Efficacy of mucosal and subcutaneous polypeptide vaccines in rhesus macaques exposed to SHIV. *J Clin Virology* (in press).

Mack, G. A. (1981). A quick and easy distribution-free test for main effects in a two-factor ANOVA, *Communic Statist Part B: Simp Comp* ***10***, 571-91.