Published in final edited form as:
Biometrics. 2009 March ; 65(1): 292–300. doi:10.1111/j.1541-0420.2008.01059.x.

Differential Equation Modeling of HIV Viral Fitness Experiments: Model Identification, Model Selection, and Multimodel Inference

Hongyu Miao[1], **Carrie Dykes**[2], **Lisa M. Demeter**[2], and **Hulin Wu**[1,*]

[1]Department of Biostatistics and Computational Biology, University of Rochester School of Medicine and Dentistry, 601 Elmwood Avenue, Box 630, Rochester, New York 14642, U.S.A

[2]Department of Medicine, University of Rochester School of Medicine and Dentistry, 601 Elmwood Avenue, Box 689, Rochester, New York 14642, U.S.A

Summary

Many biological processes and systems can be described by a set of differential equation (DE) models. However, literature in statistical inference for DE models is very sparse. We propose statistical estimation, model selection, and multimodel averaging methods for HIV viral fitness experiments in vitro that can be described by a set of nonlinear ordinary differential equations (ODE). The parameter identifiability of the ODE models is also addressed. We apply the proposed methods and techniques to experimental data of viral fitness for HIV-1 mutant 103N. We expect that the proposed modeling and inference approaches for the DE models can be widely used for a variety of biomedical studies.

Keywords

Differential equation modeling; Global optimization; HIV dual infection; Identifiability analysis; Model selection; Multimodel inference; Viral fitness

1. Introduction

Viral fitness is the ability of a pool of different viruses to replicate in a given surrounding environment. More specifically, different variants of HIV-1 viruses will compete with each other for limited resources (e.g., targeted cells) in the same environment to survive. A better understanding of HIV-1 viral fitness in vitro and in vivo is necessary for treatment strategy design such that the dominant virus during a time period can be identified and specific drugs can be selected to target this virus.

A number of independent studies have investigated HIV-1 viral fitness in vivo (Goudsmit et al., 1997) or in vitro (Martinez-Picado et al., 1999). HIV-1 replication fitness assays may use very different methods (e.g., growth competition versus parallel infection, multiple cycle versus single cycle); however, all of them compare the mutant (or “test”) virus to a reference strain (“wildtype” strain, usually drug sensitive). Although competition growth assays in vitro are usually more labor intensive, they are subject to less experimental variability and are more sensitive to subtle differences in fitness compared to parallel assays. In this article, we develop statistical methods and modeling techniques for the multiple cycle, recombinant-virus, growth

* hwu@bst.rochester.edu.

Supplementary Materials: Web Appendices, Tables, and Figures referenced in Sections 1 −3, 5, and 6 are available under the Paper Information link at the *Biometrics* website http://www.biometrics.tibs.org.

competition assay developed by Dykeset al. (2006). At hours 70, 94, 115, 139, and 163, the numbers of uninfected cells, cells infected by mutant virus, cells infected by wildtype virus, and cells infected by both viruses were calculated from the total number of cells and the percentages of uninfected cells and infected cells with different type of infections. Three replicates of measurements at each time point were obtained. For details of the experimental data, the reader is referred to Miao et al. (2008). For convenience, the data set is also listed in Web Table 1.

The development of mathematical models and statistical methods is critical to explore viral fitness assays and estimate viral fitness parameters. The relative growth kinetics of two virus variants has been investigated (Goudsmit et al., 1997; Martinez-Picado et al., 1999). In these previous studies, the measure of fitness is obtained by estimating the linear slope of the plots of the ratio of the two competing variants on a logarithmic scale against time. New definitions of relative fitness (RF) were developed by Goudsmit et al. (1997), Marée et al. (2000), and Bonhoeffer, Barbour, and De Boer (2002) by considering the ratio of production rates. However, such new definitions of RF are not consistent with the conventional definition of RF in population genetics (Wu et al., 2006). Our previous work (Wu et al., 2006) investigated the RF under the framework of viral dynamic (differential equation) models based on the assumptions of constant target (uninfected) cell concentration during the experiment and absence of dual infections, which may not be true for many experiments both in vivo (Jung et al., 2002; Levy et al., 2004) and in vitro (Dang et al., 2004; Levy et al., 2005). Miao et al. (2008) extended the mathematical model of Wu et al. (2006) to account for both dual infection and time-varying uninfected cell concentration. A time-varying viral fitness was also proposed by Miao et al. (2008) based on the viral dynamic model. The time-varying viral fitness definition is more flexible than the constant viral fitness because it can track changes in RF at different time points and can help us to better understand competition between different virus variants.

Statistical literature on inference for nonlinear ordinary differential equation (ODE) models is very sparse. For the HIV viral fitness model (nonlinear ODEs), our objective is to estimate the kinetic parameters and the corresponding relative viral fitness (a function of kinetic parameters). However, before we develop any statistical estimation methods, the identifiability of model parameters needs to be carefully studied, i.e, we need to evaluate whether all model parameters are uniquely identifiable based on the measurement data even without measurement error. Based on our experience, some of the parameters or dynamic terms in the differential equation models may be relatively small compared to other parameters or terms in the same equation. In this case, these parameters and terms may not be sensitive to the measured variables. Thus, these parameters cannot be reliably estimated from the experimental data. In this case, model selection or hypothesis testing methods may be employed to verify the importance of these parameters or terms. In addition, the multimodel inference approach, such as the model averaging procedure, may also be used for better model prediction and parameter estimation.

In this article, we will introduce the viral fitness model in Section 2. In Section 3, theoretical analysis techniques from mathematics and engineering fields (Ljung and Glad, 1994; Xia and Moog, 2003; Jeffry et al., 2005) will be used to analyze the identifiability of HIV viral fitness model. In Section 4, we employ the nonlinear least squares (NLS) method to estimate the kinetic parameters in the proposed nonlinear differential equation models; in particular, we introduce the differential evolution (DE), a global optimization algorithm to solve the NLS problem to avoid the local minima and convergence difficulty in nonlinear optimization. We suggest model selection and multimodel inference approaches in Section 5. Numerical results including experimental data analysis and limited simulation studies are reported in Section 6. Final remarks are given in Section 7.

2. HIV Viral Fitness Model with Presence of Dual Infections

For both the recombinant-virus assay and the whole-virus assay, we cannot completely rule out the presence of dually infected cells. Sometimes the dual infection rate is very small and we can ignore it. But in some other cases it may be too large to simply ignore it. Thus, we have developed a mathematical model for the situation of nonignorable dual infection, as specified as in Miao et al. (2008). The reader can also find the model details in Web Appendix A.

If we assume that the viral population densities are proportional to their corresponding infected cell densities, the model can be simplified as (Miao et al., 2008)

$$\begin{aligned}
\frac{dT}{dt} &= (\rho - k_m T_m - k_w T_w - k_R T_{mw})T \\
\frac{dT_m}{dt} &= (\rho_m + k_m T - q_m T_w)T_m + 0.25 k_R T_{mw} T \\
\frac{dT_w}{dt} &= (\rho_w + k_w T - q_w T_m)T_w + 0.25 k_R T_{mw} T \\
\frac{dT_{mw}}{dt} &= (\rho_{mw} + 0.5 k_R T)T_{mw} + (q_m + q_w)T_w T_m,
\end{aligned} \tag{1}$$

where T, T_m, T_w, and T_{mw} are numbers of uninfected cells, cells infected by mutant virus, cells infected by wildtype virus, and cells infected by both mutant and wildtype viruses (dual infection); (ρ, ρ_m, ρ_w, ρ_{mw}) the net growth rate of T, T_m, T_w, and T_{mw}, respectively; (k_m, k_w, k_R) the infection rates of mutant virus, wildtype virus, and virus from dually infected cells, respectively; and q_m and q_w the dual infection rates. Based on this model, the RF of HIV virus can be time varying. For mutant versus wildtype virus, the log-relative fitness (LRF) is defined as

$$\begin{aligned}
d_{mw}(t) = \; & g_m - g_w = (\rho_m - \rho_w) + (k_m - k_w)T(t) \\
& + q_w T_m(t) - q_m T_w(t),
\end{aligned} \tag{2}$$

where g_m and g_w are the net reproduction rates of mutant and wildtype virus, respectively. The RF of mutant versus the wildtype virus is given as

$$\begin{aligned}
\mathrm{RF}_{mw}(t) &= 1 + s = \exp\{d(t)\} \\
&= \exp\{(\rho_m - \rho_w) + (k_m - k_w)T(t) + q_w T_m(t) \\
&\quad - q_m T_w(t)\}.
\end{aligned} \tag{3}$$

Similarly, for heterozygous versus wildtype virus, the LRF and the RF are defined as

$$\begin{aligned}
d_{hw}(t) = \; & g_{mw} - g_w = \{\rho_{mw} + 0.5 k_R T(t)\} \\
& - \{\rho_w + k_w T(t) + k_R T(T) - q_w T_m(t)\},
\end{aligned} \tag{4}$$

$$\begin{aligned}
\mathrm{RF}_{hw}(t) &= 1 + s = \exp\{d(t)\} \\
&= \exp(g_{mw} - g_w),
\end{aligned} \tag{5}$$

where g_{mw} is the net reproduction rate of heterozygous viruses.

3. Theoretical Identifiability Analysis

It is critical to study the parameter identifiability of ODE models before we develop any statistical parameter estimation methods for the ODE model. We address two types of identifiability here: one is theoretical (mathematical) identifiability and another is practical (statistical) identifiability. Theoretical identifiability is to study whether the parameters in the ODE models are identifiable under the assumption of perfect measurement (no measurement error), whereas practical identifiability is to investigate whether the parameters in the ODE models are identifiable (can be reasonably estimated) with noisy experimental data. Identifiability issues for differential equation models have been investigated by engineers and mathematicians (Tunali and Tarn, 1987; Diop and Fliess, 1991; Ljung and Glad, 1994; Glad, 1997; Conte, Moog, and Perdon, 1999), and some investigators have applied these techniques to HIV dynamic models (Xia, 2003; Xia and Moog, 2003; Jeffrey and Xia, 2005). Recently we have also explored the identifiability technique for HIV models (Miao et al., 2008; Wu et al., 2008). Here we introduce the identifiability analysis in detail.

A general dynamic system in the form of differential equations can be written as:

$$\dot{x} = f(x, u, \theta), \quad y = h(x, u, \theta), \tag{6}$$

where $\boldsymbol{x} \in R^n$ is the state variable vector, $\boldsymbol{u} \in R^m$ the known input vector, $\boldsymbol{y} \in R^p$ the output vector (or measurement variables), $\boldsymbol{\theta} \in R^q$ the parameter vector, and $f \in R^n$ and $h \in R^p$ are known functions of x, u, and θ. For example in equation (A.1) in Web Appendix A, $x = (T, T_m, T_w, T_{mw}, M, W, R)$, $u = 0$, and $y = (T, T_m, T_w, T_{mw})$, f are the right-hand-side functions of all seven equations in (A.1) while h are the right-hand-side functions of the first four equations in (A.1). Identifiability analysis can answer the questions of whether θ can be uniquely identified from a given system input u and corresponding system output y and how many experimental measurements are necessary to identify θ. Such an analysis completely depends on characteristics of the system structure, so it is also called structural identifiability analysis. Structural identifiability assumes that the model structure is absolutely accurate and that the measurements are precise (no measurement error). Although these assumptions are not realistic in practice, the structural identifiability analysis is still useful. If a system is theoretically identifiable, then further practical identifiability studies can be explored via Monte Carlo simulations (Miao et al., 2008) or the correlation matrix approach (Rodriguez-Fernandez, Egea, and Banga, 2006). In this section, we focus on the structural identifiability analysis of the ODE model (1).

Linear dynamic systems are relatively simple and the corresponding identifiability problems have been well addressed, in particular for compartmental models (Audoly et al., 1998). However, for nonlinear ODE models, no general method is available to uniformly solve the identifiability problem of all nonlinear systems. Pohjanpalo (1978) proposed a method based on power series expansion of system output and Vajda, Godfrey, and Rabitz (1989) proposed a similarity transformation approach. However, these techniques are either difficult to use in practice or not suitable for high-dimensional nonlinear ODE models. Approaches based on differential algebra have been first proposed by Ljung and Glad (1994) and Ollivier (1990), and have been successfully applied to low-dimensional problems (Ljung and Glad, 1994; Audoly et al., 2001). For basic theories of differential algebra, the reader is referred to Ritt (1950). However, the differential algebra method requires that the dynamic system can be expressed in terms of differential polynomials and complicated symbolic computation algorithms are needed. General algorithms and procedures for this approach are not well established.

Xia and Moog (2003) proposed a simple alternative method based on the implicit function theorem, which has been successfully applied to HIV dynamic models with up to six dimensions (Xia and Moog, 2003; Jeffrey and Xia, 2005). This method eliminates all unobservable variables by taking derivatives of known system inputs and measurable system outputs. A number of equations consisting of unknown parameters, known system inputs and outputs, and their derivatives are obtained. Finally, the identifiability of the dynamic system can be evaluated from the derived equations. This method is practically simple but not suitable for high-dimensional nonlinear ODE models. In this study, model (1) is only four-dimensional and the procedure described in Xia and Moog (2003) and Jeffrey and Xia (2005) can be employed.

By taking derivatives of the four equations in model (1), the structural identifiability of the model can be verified. For details, the reader is referred to Web Appendix B. We found that all parameters (ρ, ρ_m, ρ_w, ρ_{mw}, k_m, k_w, k_R, q_m, q_w) are structurally identifiable if at least five measurements of T and four measurements of T_m, T_w, and T_{mw} are available under the assumption that all coefficients matrices in (B.7)–(B.11) in Web Appendix B are of full rank for at least some local time points. Note that theoretical (structural) identifiability does not imply that all the parameters in the model can be reasonably identified based on data measured with error. We will further study the practical identifiability of model (1) via Monte Carlo simulations in Section 6.

4. Model Fitting Procedure and Global Optimization

For model (1), we have measurements for each of the variables, i.e.,

$$Y_1(t_i)=T(t_i)+\varepsilon_1(t_i),\ Y_2(t_i)=T_m(t_i)+\varepsilon_2(t_i),$$
$$Y_3(t_i)=T_w(t_i)+\varepsilon_3(t_i),\ Y_4(t_i)=T_{mw}(t_i)+\varepsilon_4(t_i),$$

where $i = 1, 2,\ldots, 5$ and $\{\varepsilon_1(t_i), \varepsilon_2(t_i), \varepsilon_3(t_i), \varepsilon_4(t_i)\}$ are assumed to be independent with mean zero and common variance σ^2. Note that $T(t_i)$, $T_m(t_i)$, $T_w(t_i)$, and $T_{mw}(t_i)$ are solutions to the ODE model (1) at time t_i, and $Y_1(t_i)$, $Y_2(t_i)$, $Y_3(t_i)$, and $Y_4(t_i)$ are their corresponding measurements. The least squares (LS) method is used to minimize the LS objective function,

$$J=\sum_{i=1}^{n}\Big[\omega_T\{Y_1(t_i)-T(t_i,\theta)\}^2+\omega_m\{Y_2(t_i)-T_m(t_i,\theta)\}^2 +\omega_w\{Y_3(t_i)-T_w(t_i,\theta)\}^2+\omega_{mw}\{Y_4(t_i)-T_{mw}(t_i,\theta)\}^2\Big],$$

where $T(t_i, \theta)$, $T_m(t_i, \theta)$, $T_w(t_i, \theta)$, and $T_{mw}(t_i, \theta)$ need to be evaluated numerically from the ODE model (1). We weight the four measured variables by ω_T, ω_m, ω_w, and ω_{mw} using simple weight strategy of equal weight, i.e., $\omega_T = \omega_m = \omega_w = \omega_{mw} = 1$. Alternatively, we may also weight the four variables by considering their units (scales) and measurement errors. If the data are correlated, the generalized least squares (GLS) method may be used in this case. To minimize the above objective function, we need to use a nonlinear optimization algorithm. The bootstrap method (Shao and Tu, 1995) for nonlinear regression models is used to obtain the confidence intervals.

There are mainly two different categories of optimization algorithms: gradient and global methods. Gradient methods, such as the Levenberg–Marquardt method and the Gauss–Newton method, usually require an initial starting point in the parameter space that is close to the true solution point, then the steepest descent direction is calculated based on the first- and/or second-

order local gradient of the objective function in the parameter space to determine the search direction for the next step, and so on. The gradient methods are computationally efficient. However, if the objective function is not smooth enough or even discrete, these gradient methods may not work. In addition, for multimodal objective functions, the search direction calculated from the local gradient is very likely to be misleading. Under such situations, the gradient methods are very sensitive to the position of the starting values of the unknown parameters and they may be trapped by local minima. For details of gradient methods and their applications in ODE parameter estimation, the reader is referred to Nocedal and Wright (1999) and Englezos and Kalogerakis (2001). Although global optimization methods are usually less computationally efficient compared to gradient methods, global optimization methods are more flexible being able to handle multimodal, nonsmooth, and even discrete objective functions. For instance, Linga, Al-Saifi, and Englezos (2006) compared the Luus–Jaakola method (Luus and Jaakola, 1973), one of the global optimization methods, to the Gauss–Newton method for parameter estimation of ODE models and they found that the two methods yielded very different results for the same problem, but the sum of squared residuals achieved by the Luus–Jaakola method was much smaller than that from the Gauss–Newton method. In addition, global optimization methods usually do not require an initial starting point; instead only a search range of parameters needs to be specified.

A number of global optimization methods have been developed. The performance of global optimization methods can be evaluated by computational efficiency and the probability of being trapped by local minima. Seven global optimization algorithms, including the Luus–Jaakola method (Luus and Jaakola, 1973) and the DE method (Storn and Price, 1997), were evaluated by Moles, Bangaa, and Keller (2004). Compared to six other methods, the DE method has a lower probability of being trapped and a moderate computational cost. We also compared the performance of the DE method to several gradient methods such as the Gauss–Newton method, the Levenberg–Marquardt method, and the quasi-Newton method via simulations (data not shown). It was found that for model (1), the gradient method failed to locate the minima of the objective function even if the starting point was just 5% away from the true parameter values. Thus we selected to use the DE method in this article.

There are four stages in the DE method: initialization, mutation, crossover, and selection (Storn and Price, 1997). The initial population is uniformly and randomly generated within the search range to attempt to cover the whole region, denoted by $x_{i,G}(i = 1, 2,\ldots, NP)$, where NP is the number of members in this generation G. Then mutated populations are generated by randomly mixing the previous generation with certain weights. Storn and Price (1997) suggested a mutation method $v_{i,G+1} = x_{r1,G} + F \cdot (x_{r2,G} - x_{r3,G})$, where $v_{i,G+1}$ is the ith member of generation $(G + 1)$, $x_{r1,G}$, $x_{r2,G}$, and $x_{r3,G}$ the members of the previous generation G, and $F > 0$ the amplification factor. Integers r_1, r_2, and r_3 are randomly chosen from $\{1, 2,\ldots, NP\}$, which are mutually different from each other and different from i. Also, to increase diversity, the mutated member $v_{i,G+1}$ exchanges its components with $x_{i,G}$ with a given probability, the crossover ratio. For this purpose, a number within [0, 1] is generated for each component of $v_{i,G+1}$ by a uniform random number generator. If this number is greater than the crossover ratio, then the component of $x_{i,G}$ is kept. Finally, the best member in the mutated population is selected by comparing the values of objective function. The DE method is one of the genetic algorithms and its convergence rate depends on specific mutation strategies used, i.e., the amplification factor and the crossover ratio. For more details about the DE algorithm, the reader is referred to Storn and Price (1997).

5. Model Selection and Multimodel Inference

Model selection and evaluation are important in biomedical research because many variations of models are available under different biological assumptions. To test which biological

assumption is more plausible to fit the experimental data, we need to employ model selection methods to evaluate different models and multimodel inference may also be considered, if necessary. For our HIV viral fitness experiment, model (1) is considered as the full model, based on which a number of submodels are derived under different assumptions. We list these submodels in Table 1. These assumptions are based on either biological hypothesis or preliminary data analysis. In this section, we apply the model selection method and the multimodel inference technique to the HIV viral fitness model (1).

5.1 Model Selection and Comparison

Standard model selection criteria include Akaike information criterion (AIC) (Akaike, 1973) and the Schwarz criterion (SC), also known as the Bayesian information criterion (BIC; Schwarz, 1978), and their variations such as AICc, which is the AIC with a bias correction term for small sample size. Under a likelihood framework, these criteria can be written as

$$\begin{aligned} &\mathrm{AIC} = -2\mathrm{In}L + 2K, \mathrm{BIC} = -2\mathrm{In}L + K\mathrm{In}(n), \\ &\mathrm{AICc} = \mathrm{AIC} + 2K(K+1)/(n-K-1), \end{aligned} \tag{7}$$

where L is the likelihood function, K the number of parameters, and n the sample size. If the model error can be assumed to be independent and normally distributed, the above criteria can be written as

$$\begin{aligned} &\mathrm{AIC} = n\mathrm{In}\left(\frac{\mathrm{RSS}}{n}\right) + 2K, \mathrm{BIC} = n\mathrm{In}\left(\frac{\mathrm{RSS}}{n}\right) + K\mathrm{In}(n), \\ &\mathrm{AICc} = n\mathrm{In}\left(\frac{\mathrm{RSS}}{n}\right) + \frac{2nK}{n-K-1}, \end{aligned} \tag{8}$$

where RSS denotes the residual sum of squares. Note that the AIC and AICc are different due to the difference in the term of model complexity penalty. As a practical guideline, if the number of unknown parameters exceeds $n/40$ (where n is the sample size), the AICc instead of AIC should be used. For our experiment, the sample size n is equal to 60, and the number of unknown parameters varies between 8 and 13 for different submodels (unknown initial values are also included), which is much larger than $n/40 = 60/40 = 1.5$. So the AICc is more appropriate. In fact, the AICc converges to AIC as the sample size gets larger, thus the AICc is often suggested to be employed regardless of the sample size (Burnham and Anderson, 2004). For comparison, we calculated AIC, AICc, and BIC. In addition, the likelihood ratio test (Mood, Graybill, and Boes, 1974; Bickel and Doksum, 2001) was also used to evaluate whether the difference between two models is statistically significant. For details of likelihood ratio test, please refer to Web Appendix C.

We calculated the AIC, BIC, and AICc based on the formulas (8) numerically for fitting model (1). To obtain the RSS, the fitted observations can be calculated by evaluating the ODE model (1) numerically based on the estimated parameters. The likelihood function can also be obtained based on the RSS under the normal assumption of the measurement errors. This allows us to calculate the test statistics of the likelihood ratio test.

5.2 Model Averaging

Inference based on the best model that is selected from a model selection method does not account for model uncertainty. Because models, especially biomedical models, are just an approximation of reality, model uncertainty actually always exists and therefore should be considered when multiple reasonable models are available. Multimodel inference may be

necessary in some cases. Multimodel inference not only accounts for model uncertainty, but also is important when several models are equally well supported by the experimental data. Burnham and Anderson (2004) compared the performance of the model averaging strategy to that of the best model strategy by considering the predictive mean square error based on Monte Carlo simulations. They concluded from their work and the literature that the model averaging strategy is generally better than the best model strategy for predictions.

In this project, model weights based on model selection criteria were calculated to account for model uncertainty and the weighted average of parameter values was calculated. The AICc was used for our data due to the small sample size. Burnham and Anderson (2004) discussed the difference between AICc and BIC, and compared their performance under different scenarios. The selected model based on BIC was found to be underfit given realistic sample sizes. In the tapering-effects scenario, the AICc outperforms BIC (Burnham and Anderson, 2004). In our data analysis, the model averaging strategy was based on AICc weights. A brief introduction to model averaging based on Akaike weights is given here and more details can be found in Burnham and Anderson (2004).

Assume that R models are available and the individual AICc of each model is obtained, then define the minimum AICc value among all R models as $AICc_{min}$ and define the difference

$$\Delta_i = AICc_i - AICc_{\min}, i = 1, 2, 3, \ldots, R. \tag{9}$$

Then the Akaike weight of each model is defined as

$$w_i = \exp(-\Delta_i/2) \Big/ \sum_{r=1}^{R} \exp(-\Delta_r/2), \quad i = 1, 2, 3, \ldots, R. \tag{10}$$

Obviously, all the Akaike weights sum to one and the value of Akaike weights depend on both the individual model and the total number of models. The estimators of the parameters and variances based on Akaike weights are

$$\bar{\hat{\theta}} = \sum_{i=1}^{R} w_i \hat{\theta}_i, \tag{11}$$

$$\operatorname{var}(\bar{\hat{\theta}}) = \left[\sum_{i=1}^{R} w_i \left\{ \operatorname{var}(\hat{\theta}_i | g_i) + (\hat{\theta}_i - \bar{\hat{\theta}})^2 \right\}^{1/2} \right]^2, \tag{12}$$

where $\hat{\theta}_i$ is the estimated parameter vector of model g_i.

6. Numerical Results: Relative Viral Fitness of HIV Mutant 103N

First we report the data analysis results for the RF of HIV mutant 103N (Miao et al., 2008, see also Web Table 1) based on the model selection method. We fitted all submodels (Table 1)

and the full model to the experimental data of HIV-1 103N mutant and wildtype viruses using the LS method and the proposed algorithm in Section 4 (see Figure 1 for fitted curves). These AIC, BIC, or AICc values for all the submodels are reported in Table 2. The smaller AIC, BIC, or AICc value corresponds to a better model. From Table 2, we can see that the AIC selects submodel 2 as the best model and submodel 3 as the second-best model, while the AICc selects submodel 3 as the best model and submodel 2 as the second-best model. However, the BIC selects submodel 7 as the best model and submodel 3 as the second-best model. Thus, our following discussion and data analyses based on the selected models will focus on submodel 3 ($\rho_m = 0$, $\rho_w = 0$, $\rho_{mw} = 0$) and submodel 7 ($\rho_m = 0$, $\rho_w = 0$, $\rho_{mw} = 0$, $k_R = 0$). Note that the difference between these two models is very small, e.g., the relative differences of the BIC and AICc of the two models are ΔBIC = 0.14% and ΔAICc = 0.19%, respectively. In Table 2, the log-likelihood values ($-2 \ln L$) and the p-values from likelihood ratio test are also reported to compare all submodels to the full model. These results suggest that most submodels except submodels 11 and 15 fit the data equally as well as the full model does.

The parameter estimation results for submodels 3 and 7 are reported in Table 3. Model fitting for both models is reasonably good (Figure 1) and we present goodness of fit and residual plots in Web Figures 2–5. These residual analyses have no indication of lack of fit. Comparing the estimates from both submodels, we found that the difference in the estimates between the two submodels is small. The difference in parameters (ρ, k_m, k_w, q_w) is less than 5%, and the difference in parameter q_m is about 17%. The 95% confidence intervals of all estimates are also given in Table 3 using the bootstrap method (Shao and Tu, 1995;Davison and Hinkley, 1997). From these results, we can see that the dual infection rates (q_m, q_w) were higher than the regular infection rates (k_m, k_w) in this experiment. The net growth rate of the uninfected target cells was estimated as $\rho = 0.015$, which indicates that the doubling time of uninfected cells is 46 hours. The bootstrap confidence interval of k_R from submodel 3 is too wide (1.12E–21, 9.55E–10), which indicates that this estimate may not be reliable. Thus, submodel 7 may be better.

To evaluate whether these estimates are reliable, similar to Miao et al. (2008), we also performed Monte Carlo simulations (Metropolis and Ulam, 1949) with the assumption of different measurement error levels (5%, 10%, and 30%). We define the average relative estimation error (ARE) as $\mathrm{ARE}=1/N\sum_{j=1}^{N}|\theta-\widehat{\theta}_j|/|\theta|\times 100\%$, where $\hat{\theta}_j$ is the estimate of parameter θ from the jth simulation data set and $N = 1000$ is the total number of simulation runs. We use the ARE to evaluate the estimates of the kinetic parameters and the results are also reported in Table 3 (last three columns). From these results, we can see that the AREs of k_R in submodel 3 and q_w in submodel 7 are relatively large compared to those of other parameters. For small measurement errors with large sample size, the estimates of all the parameters in submodels 3 and 7 are reasonable. However, if the measurement error is large and the sample size is small, the parameters (q_w, k_R) may not be reliably estimated.

We also applied the model averaging method introduced in Section 5.2 to this data set and the results are reported in Table 4. We used all submodels in Table 1 for the model averaging analysis. As we expected, the estimates of some parameters may not be reliable because their confidence intervals are large and cover 0. However, the estimates of the four parameters ρ, k_m, k_w, and q_w can be reliably estimated, which is consistent with the conclusions reached from the model selection methods.

We also calculated the time-varying RF parameters, d_{mw}, RF_{mw}, d_{hw}, and RF_{hw} from the formulas (2)–(5) based on the parameter estimates from submodels 3 and 7 at all five observation points. These results are presented in Table 5. Notice that both the mutant virus and the heterozygous virus compete with the wildtype virus and also compete with each other

in the same environment. Their RF estimates versus wildtype virus are expected to be negatively related (see Table 5 and Web Figure 1). For both submodels 3 and 7, d_{mw} and RF_{mw} monotonically decrease from hour 70 to hour 163; however, d_{hw} and RF_{hw} decrease at hours 70, 94, 115, and 139, and then increase at hour 163. The nonmonotone trajectories of d_{hw} indicate that the dominance of different types of viruses in terms of fitness changes over time.

7. Discussion and Conclusion

Differential equation modeling is an important tool in infectious disease research such as HIV dynamics and cellular kinetics. However, very little statistical literature formally deals with statistical inference for differential equation models. In particular, it is very important to investigate the identifiability of ODE models before we develop parameter estimation methods. Because the parameter estimation often requires numerical evaluations of differential equations, the computational cost is high and the computational algorithm may easily run into convergence problems. Model selection and multimodel inference need to be considered.

In this article, we introduced a complete procedure and related techniques for modeling HIV viral fitness experiments using differential equations. We considered theoretical and practical identifiability, the model fitting procedure, model selection, and multimodel inference techniques, which are spread across the engineering, mathematics, and statistics literature, for differential equation models. To model a biological experiment or process such as the HIV viral fitness experiment that was considered here, we suggested a modeling and data analysis procedure consisting of several steps: (i) develop mathematical models based on the biological mechanism of the experiment; (ii) perform the identifiability analysis of the derived models based on the experimental measurements; (iii) develop parameter estimation and inference methods for the proposed models; (iv) perform model evaluation, selection or multimodel inference; and (v) evaluate model and diagnose goodness of fit. We have illustrated this modeling and data analysis strategy using the HIV viral fitness experiment.

Based on our previously developed full model for HIV viral fitness experiments (Miao et al., 2008), we derived a number of submodels based on biological assumptions and preliminary data exploration. The model selection methods such as AIC, BIC, and AICc were employed to evaluate these submodels. As we expected and similar to statistical regression models, the best ODE model selected by BIC has fewer parameters than the best models selected by AIC and AICc. The likelihood ratio test was also used to assess statistically significant differences between the two ODE models. The evaluation analysis suggested to us to focus on submodels 3 and 7 for our experimental data. In addition, we have also explored the model averaging method to obtain the estimates for more biological parameters compared to that in the selected best models. We applied the proposed models and methods to an experimental data set for studying the RF of HIV mutant 103N versus wildtype virus. HIV viral fitness is calculated based on the two best models. The patterns of d_{hw} and RF_{hw} in models 3 and 7 are found to be similar to the patterns of d_{hw} and RF_{hw} in the full model (Miao et al., 2008), but the patterns of d_{mw} and RF_{mw} in the two submodels are different from those in the full model for a different HIV mutant virus.

In summary, we have introduced some basic statistical inference methods for differential equation models, which include parameter estimation, model selection, model averaging, and hypothesis testing. We have also studied the identifiability of differential equation models, which is relatively new to the statistical research community, but very useful and important when applying ODE models to biomedical research. We expect that the proposed modeling and statistical approaches could provide general guidance for differential equation modeling of biomedical systems, and at the same time, stimulate more statistical research in this direction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the NIAID/NIH grants AI50020, AI052765, AI055290, AI 065217, and AI27658.

References

Akaike, H. Information theory as an extension of the maximum likelihood principle. In: Petrov, BN.; Csaki, F., editors. Second International Symposiumon Information Theory. Budapest: Akademiai Kiado; 1973. p. 267-281.

Audoly S, D'angio L, Saccomani MP, Cobelli C. Global identifiability of linear compartmental models. IEEE Transactions on Biomedical Engineering 1998;45:36–47. [PubMed: 9444838]

Audoly S, Bellu G, D'Angio L, Saccomani MP, Cobelli C. Global identifiability of nonlinear models of biological systems. IEEE Transactions on Biomedical Engineering 2001;48:55–65. [PubMed: 11235592]

Bickel, PJ.; Doksum, KA. Mathematical Statistics Basic Ideas and Selected Topics. 2nd. Englewood Cliffs, NJ: Prentice Hall; 2001.

Bonhoeffer S, Barbour AD, De Boer RJ. Procedures for reliable estimation of viral fitness from time-series data. Proceedings of the Royal Society of London 2002;269:1887–1893.B

Burnham KP, Anderson DR. Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods Research 2004;33:261–304.

Conte, G.; Moog, CH.; Perdon, AM. Nonlinear Control Systems: An Algebraic Setting. London: Springer; 1999.

Dang Q, Chen J, Unutmaz D, Coffin JM, Pathak VK, Powell D, KewalRamani VN, Maldarelli F, Hu WS. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proceedings of the National Academy of Sciences USA 2004;101:632–637.

Davison, AC.; Hinkley, DV. Bootstrap Methods and Their Application. Cambridge, U.K.: Cambridge University Press; 1997.

Diop, S.; Fliess, M. On nonlinear observability. Proceedings of the First European Control Conference; Paris, Hermes. 1991. p. 152-157.

Dykes C, Wang J, Jin X, Planelles V, An DS, Tallo A, Huang Y, Wu H, Demeter LM. Evaluation of a multiple-cycle, recombinant virus, growth competition assay that uses flow cytometry to measure replication efficiency of HIV-1 in cell culture. Journal of Clinical Microbiology 2006;44:1930–1943. [PubMed: 16757582]

Englezos, P.; Kalogerakis, N. Applied Parameter Estimation for Chemical Engineers. New York: Marcel-Dekker; 2001.

Goudsmit J, De Ronde A, De Rooij E, De Boer RJ. Broad spectrum of in vivo fitness of human immunodeficiency virus type 1 subpopulations differing at reverse transcriptase codons 41 and 215. Journal of Virology 1997;71:4479–4484. [PubMed: 9151839]

Glad, ST. Solvability of differential algebraic equations and inequalities: An algorithm. European Control Conference; Brussels. 1997. ECC97

Jeffrey, AM.; Xia, X. Identifiability of HIV/AIDS models. In: Tan, WY.; Wu, H., editors. Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention. Singapore: World Scientific; 2005. p. 255-286.

Jung A, Maier R, Vartanian JP, Bocharov G, Jung V, Fischer U, Meese E, Wain-Hobson S, Mayerhans A. Multiply infected spleen cells in HIV patients. Nature 2002;418:144. [PubMed: 12110879]

Levy DN, Aldrovandi GM, Kutsch O, Shaw GM. Dynamics of HIV-1 recombination in its natural target cells. Proceedings of the National Academy of Sciences USA 2005;101:4204–4209. and correction (2005) **102**, 1808.

Linga P, Al-Saifi N, Englezos P. Comparison of the Luus-Jaakola optimization and Gauss-Newton methods for parameter estimation in ordinary differential equation models. Industrial & Engineering Chemistry Research 2006;45:4716–4725.

Ljung L, Glad ST. On global identifiability for arbitrary model parameterizations. Automatica 1994;30:265–276.

Luus R, Jaakola THI. Optimization by direct search and systematic reduction of size of search region. AIChE Journal 1973;19:760–766.

Martinez-Picado J, Savara AV, Sutton L, D'Aquila RT. Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. Journal of Virology 1999;73:3744–3752. [PubMed: 10196268]

Marée AFM, Keulen W, Boucher CAB, De Boer RJ. Estimating relative fitness in viral competition experiments. Journal of Virology 2000;74:11067–11072. [PubMed: 11070001]

Metropolis N, Ulam S. The Monte Carlo method. Journal of the American Statistical Association 1949;44:335–341. [PubMed: 18139350]

Miao H, Dykes C, Demeter LM, Cavenaugha J, Parka SY, Perelson AS, Wu H. Modeling and estimation of kinetic parameters and replication fitness of HIV-1 from flow-cytometry-based growth competition experiments. Bulletin of Mathematical Biology. 2008

Moles CG, Bangaa JR, Keller K. Solving nonconvex climate control problems: Pitfalls and algorithm performances. Applied Soft Computing 2004;5:35–44.

Mood, AM.; Graybill, FA.; Boes, DC. Introduction to the Theory of Statistics. 3rd. New York: McGraw-Hill; 1974.

Nocedal, J.; Wright, SJ. Numerical Optimization. New York: Springer Verlag; 1999.

Ollivier, F. Thèsis de Doctorat en Science. École Polytéchnique; Paris, France: 1990. Le problème de l'identifiabilité structurelle glob-ale: étude théorique, méthodes effectives et bornes de complexité.

Pohjanpalo H. System identifiability based on the power series expansion of the solution. Mathematical Biosciences 1978;41:1–32.

Ritt, J. Differential Algebra. Providence, RI: American Mathematical Society; 1950.

Rodriguez-Fernandez M, Egea JA, Banga JR. Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. BMC Bioinformatics 2006;7:483–500. [PubMed: 17081289]

Schwarz G. Estimating the dimensions of a model. Annals of Statistics 1978;6:461–464.

Shao, J.; Tu, D. The Jackknife and Bootstrap. New York: Springer; 1995.

Storn R, Price K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 1997;11:341–359.

Tunali T, Tarn TJ. New results for identifiability of nonlinear systems. IEEE Transactions on Automatic Control 1987;32:146–154.

Vajda S, Godfrey K, Rabitz H. Similarity transformation approach to identifiability analysis of nonlinear compartmental models. Mathematical Biosciences 1989;93:217–248. [PubMed: 2520030]

Wu H, Huang Y, Dykes C, Liu D, Ma J, Perelson AS, Demeter L. Modeling and estimation of replication fitness of HIV-1 in vitro experiments using a growth competition assay. Journal of Virology 2006;80:2380–2389. [PubMed: 16474144]

Wu H, Zhu H, Miao H, Perelson AS. Parameter identifiability and estimation of HIV/AIDS dynamic models. Bulletin of Mathematical Biology 2008;70:785–799. [PubMed: 18247096]

Xia X. Estimation of HIV/AIDS parameters. Automatica 2003;39:1983–1988.

Xia X, Moog CH. Identifiability of nonlinear systems with applications to HIV/AIDS models. IEEE Transactions on Automatic Control 2003;48:330–336.

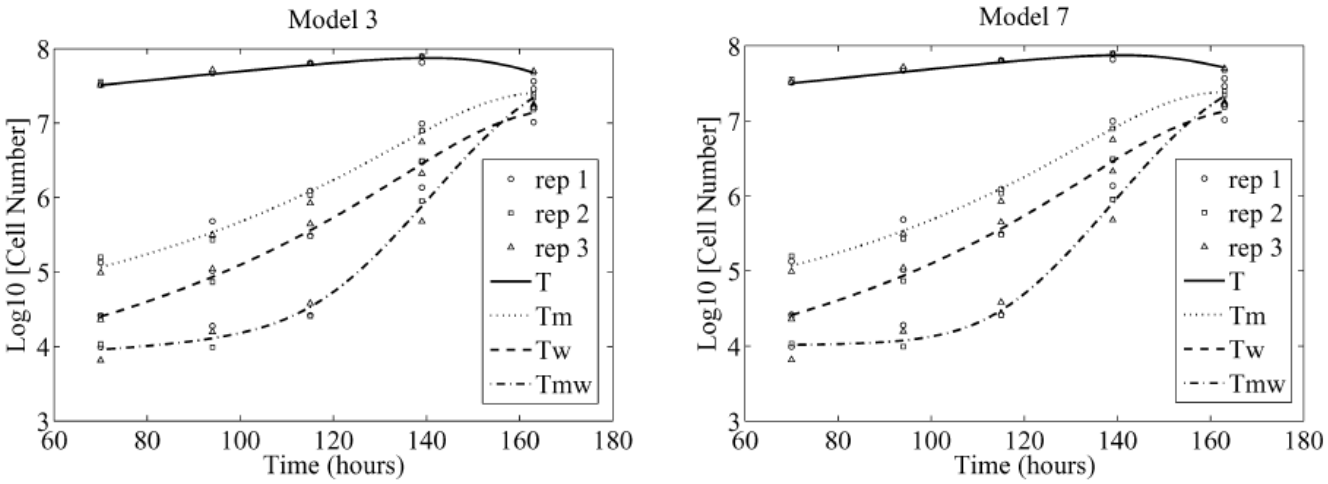

Figure 1.
Model fitting to HIV mutant 103N experimental data using both models 3 and 7.

Table 1
Submodels and corresponding assumptions

Submodel	Assumptions
1	$\rho_{mw} = 0$
2	$\rho_w = 0, \rho_{mw} = 0$
3	$\rho_m = 0, \rho_w = 0, \rho_{mw} = 0$
4	$k_R = 0$
5	$\rho_{mw} = 0, k_R = 0$
6	$\rho_w = 0, \rho_{mw} = 0, k_R = 0$
7	$\rho_m = 0, \rho_w = 0, \rho_{mw} = 0, k_R = 0$
8	$q_w = 0$
9	$\rho_{mw} = 0, q_w = 0$
10	$\rho_w = 0, \rho_{mw} = 0, q_w = 0$
11	$\rho_m = 0, \rho_w = 0, \rho_{mw} = 0, q_w = 0$
12	$k_R = 0, q_w = 0$
13	$\rho_{mw} = 0, k_R = 0, q_w = 0$
14	$\rho_w = 0, \rho_{mw} = 0, k_R = 0, q_w = 0$
15	$\rho_m = 0, \rho_w = 0, \rho_{mw} = 0, k_R = 0, q_w = 0$
16	$k_m = k_w$
17	$q_m = q_w$
18	$k_m = k_w, q_m = q_w$

Table 2

Model selection for the experimental data of HIV 103N mutant

Model	AIC	BIC	AICc	Log likelihood (−2 ln *L*)	*p*-value
1	−1.625E+02	−1.374E+02	−1.565E+02	−1.865E+02	0.9773
2	−1.644E+02	−1.413E+02	−1.594E+02	−1.864E+02	0.9290
3	−1.637E+02	−1.428E+02	−1.597E+02	−1.837E+02	0.4241
4	−1.622E+02	−1.371E+02	−1.562E+02	−1.862E+02	0.5894
5	−1.610E+02	−1.379E+02	−1.560E+02	−1.830E+02	0.1689
6	−1.624E+02	−1.414E+02	−1.584E+02	−1.824E+02	0.2430
7	−1.619E+02	−1.430E+02	−1.589E+02	−1.799E+02	0.1565
8	−1.600E+02	−1.349E+02	−1.540E+02	−1.840E+02	0.1142
9	−1.612E+02	−1.381E+02	−1.562E+02	−1.832E+02	0.1865
10	−1.604E+02	−1.395E+02	−1.564E+02	−1.804E+02	0.1072
11	−1.581E+02	−1.392E+02	−1.551E+02	−1.761E+02	0.0332
12	−1.620E+02	−1.390E+02	−1.570E+02	−1.840E+02	0.2873
13	−1.611E+02	−1.402E+02	−1.571E+02	−1.811E+02	0.1451
14	−1.594E+02	−1.405E+02	−1.564E+02	−1.774E+02	0.0572
15	−1.575E+02	−1.407E+02	−1.555E+02	−1.735E+02	0.0228
16	−1.619E+02	−1.368E+02	−1.559E+02	−1.859E+02	0.4328
17	−1.592E+02	−1.341E+02	−1.532E+02	−1.832E+02	0.0678
18	−1.605E+02	−1.374E+02	−1.555E+02	−1.825E+02	0.1314
Full	−1.605E+02	−1.333E+02	−1.535E+02	−1.865E+02	—

Table 3

Model fitting to experimental data and numerical sensitivity analysis based on simulation studies for the experimental data of HIV 103N mutant. The ARE is calculated based on 1000 simulation runs; 1000 simulated replicates are used for each simulation run

				ARE (%)		
Submodel	**Parameters**	**Estimate**	**95% bootstrap confidence interval**	**σ = 5%**	**σ = 10%**	**σ = 30%**
3	ρ	1.45E–02	1.23E–02, 1.68E–02	1.06	2.38	5.48
	k_m	1.16E–09	1.05E–09, 1.27E–09	0.86	1.78	4.75
	k_w	1.31E–09	1.17E–09, 1.45E–09	1.02	2.15	4.71
	k_R	4.92E–10	1.12E–21, 9.55E–10	7.54	14.90	44.60
	q_m	3.60E–09	2.09E–09, 5.32E–09	3.27	6.84	20.80
	q_w	1.57E–09	5.00E–10, 2.81E–09	1.06	11.10	32.50
7	ρ	1.50E–02	1.27E–02, 1.73E–02	1.10	2.14	6.05
	k_m	1.20E–09	1.09E–09, 1.31E–09	0.78	1.58	4.53
	k_w	1.34E–09	1.20E–09, 1.49E–09	0.89	1.53	5.25
	q_m	4.31E–09	2.77E–09, 6.01E–09	2.86	4.31	16.90
	q_w	1.50E–09	3.47E–10, 2.76E–09	5.56	10.90	37.60

Table 4

Parameter values and standard deviations based on Akaike weights for the experimental data of HIV 103N mutant

Parameter	Value	Standard deviation
ρ (10^{-2})	1.49E–02	±1.97E–03
ρ_m (10^{-2})	−1.47E–02	±2.09E–02
ρ_w (10^{-2})	4.81E–03	±1.39E–02
ρ_{mw} (10^{-2})	4.45E–04	±4.15E–03
k_m (10^{-9})	1.40E–09	±3.37E–10
k_w (10^{-9})	1.16E–09	±3.22E–10
k_R (10^{-22})	2.80E–10	±4.44E–10
q_m (10^{-9})	4.40E–09	±1.64E–09
q_w (10^{-22})	1.03E–09	±1.12E–09

Table 5

The estimate of LRF (d) and RF for mutant virus versus wildtype virus and heterozygous virus versus wildtype virus for submodels 3 and 7 based on experimental data for HIV 103N mutant. 95% bootstrap confidence intervals are listed at the bottom of each estimate

Submodel	Time (hour)	$d_{mw}(t)$ (mutant versus wild)	$RF_{mw}(t)$ (mutant versus wild)	$d_{hw}(t)$ (heterozygous versus wild)	$RF_{hw}(t)$ (heterozygous versus wild)
3	70	−4.75E–03 (−9.56E–03, −1.61E–04)	9.95E–01 (9.90E–01, 1.00E+00)	−5.00E–02 (−5.94E–02, −4.07E–02)	9.51E–01 (9.42E–01, 9.60E–01)
	94	−6.59E–03 (−1.30E–02, −3.56E–04)	9.93E–01 (9.87E–01, 1.00E+00)	−7.01E–02 (−8.06E–02, −5.86E–02)	9.32E–01 (9.23E–01, 9.43E–01)
	115	−8.45E–03 (−8.75E–03, −1.65E–02)	9.92E–01 (9.84E–01, 9.99E–01)	−9.19E–02 (−1.05E–01, −7.78E–02)	9.12E–01 (9.00E–01, 9.25E–01)
	139	−9.73E–03 (−1.63E–02, −4.15E–03)	9.90E–01 (9.84E–01, 9.96E–01)	−1.05E–01 (−1.22E–01, −8.79E–02)	9.01E–01 (8.85E–01, 9.16E–01)
	163	−1.68E–02 (−6.49E–02, 3.27E–02)	9.83E–01 (9.37E–01, 1.03E+00)	−3.39E–02 (−5.58E–02, −5.08E–03)	9.67E–01 (9.46E–01, 9.95E–01)
	Mean SD	−9.27E–03 4.63E–03	9.91E–01 4.58E–03	−7.01E–02 2.90E–02	9.33E–01 2.71E–02
7	70	−4.33E–03 (−9.33E–03, 3.67E–04)	9.96E–01 (9.91E–01, 1.00E+00)	−4.19E–02 (−4.72E–02,−3.66E–02)	9.59E–01 (9.54E–01, 9.64E–01)
	94	−6.10E–03 (−1.28E–02, 3.43E–04)	9.94E–01 (9.87E–01, 1.00E+00)	−5.93E–02 (−6.45E–02, −5.42E–02)	9.42E–01 (9.38E–01, 9.47E–01)
	115	−8.05E–03 (−1.62E–02, −3.16E–04)	9.92E–01 (9.84E–01, 1.00E+00)	−7.82E–02 (−8.39E–02, −7.29E–02)	9.25E–01 (9.20E–01, 9.30E–01)
	139	−1.10E–02 (−1.72E–02, −5.61E–03)	9.89E–01 (9.83E–01, 9.94E–01)	−8.78E–02 (−9.42E–02, −8.18E–02)	9.16E–01 (9.10E–01, 9.21E–01)
	163	−2.83E–02 (−7.56E–02, 2.04E–02)	9.72E–01 (9.27E–01, 1.02E+00)	−3.20E–02 (−5.36E–02, −4.38E–03)	9.69E–01 (9.48E–01, 9.96E–01)
	Mean SD	−1.15E–02 9.67E–03	9.89E–01 9.50E–03	−5.99E–02 2.35E–02	9.42E–01 2.22E–02