

Dose Response Models for Multiple Endpoints: A Simulation Study

Farkad Ezzet (1), Eyas Raddad (2)

(1) Aycer Pharma Consulting, Chatham, NJ, USA, (2) Chorus, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, USA

INTRODUCTION

Clinical trials often collect data on multiple endpoints. There are potential advantages in simultaneously modeling multiple endpoints in a single model fit. Such advantages include the ability to:

- Perform tests on differences between shared parameters between endpoints, e.g. time of 50% of maximal effect, ET_{50}
- Improve model fit when an endpoint is sparsely sampled by borrowing information from other endpoints

Using simulation we aim to explore the performance of bivariate and multivariate dose response models and contrast performance with the univariate approach

OBJECTIVE

To assess the performance of bivariate and multivariate dose response models and explore accuracy of model estimates in the case of sparse data for one endpoint

METHODS

Consider $Y_{ij,k}(t)$ to be the response for k^{th} endpoint for the i^{th} subject when treated with the j^{th} dose at time t ,

$$Y_{ij,k}(t) = f_k(\theta_{ij}, t) + \{ \delta_{1k} + f_k(\theta_{ij}, t)^{\delta_{2k}} \} \varepsilon_{ij,k}(t), \quad k=1,2,\dots,M$$

The parameters δ_1 and δ_2 alongside $\varepsilon_{ij,k}$ (assumed $N(0, \sigma_k^2)$) describe an additive and power error model, with $\sigma_k = \gamma_k \cdot \sigma_1$, and γ_k being the fold difference in error standard deviation of the for k^{th} endpoint as compared to the 1st.

We initially investigated 2 endpoints (bivariate) that follow an Emax model in time, but with an inhibitory I_{max} in dose for the first endpoint and a stimulatory E_{max} in dose for the second endpoint, as follows:

$$Y_{ij,1} = T_{max,1} \left(t / (t + ET_{50,1}) \cdot [1 - I_{max} \{ Dose_{ij} / (Dose_{ij} + ID_{50}) \}] \right) \quad (1)$$

$$Y_{ij,2} = [T_{max,2} + E_{max} \{ Dose_{ij} / (Dose_{ij} + ED_{50}) \}] \left(t / (t + ET_{50,2}) \right) \quad (2)$$

Where $T_{max,i}$ is the maximal time effect for endpoint $i=1, 2$.

Simulation:

1. For the bivariate case, a data set (N=24) from a 3-period X-over was simulated (Figure 1). Subjects received 3 doses of 0.001, 1, and 10 mg. At each dose, Y_1 was observed at $t=1, 3, 6, 12$ and 24, whereas Y_2 was observed only at $t=12$ and 24. Parameter values were, $T_{max,1}=50, T_{max,2}=5, \log(ET_{50,1})=1, \log(ET_{50,2})=1.5, I_{max}=0.8, E_{max}=1, ID_{50}=ED_{50}=1$. The error model was proportional with $\delta_{11}=\delta_{12}=0, \delta_{21}=1.5, \delta_{22}=1$ and $\sigma_1=\sigma_2=0.05$ (i.e. $\gamma=1$). Only $T_{max,i}$ was assumed to vary between subjects, according to a proportional error, common to both endpoints, with CV of $IIV=0.1$. Data set is created by stacking Y_1 and Y_2 as a single Y variable with a flag identifying the two endpoints. To investigate bias from the univariate analysis of Y_2 , scenarios of increasing sample size (N) up to 72 or increasing number of time points up to 6 were considered.

2. For the multivariate case, we simulated a data set using 5 endpoints (Y_1, \dots, Y_5) such that all endpoints follow an I_{max} model (equation 1), N=24 and $t=1, 3, 6, 12$ and 24. True parameter values for $T_{max,i}$ and $I_{max,i}$ were 10, 20, 30, 40, 50 and .1, .2, .3, .4, .5, respectively. $\log(ET_{50,i})=ED_{50,i}=1, CV$ of IIV in $T_{max,i}=0.1, \delta_{1i}=0, \delta_{2i}=1$ and $\sigma_i=0.05, i=1, \dots, 5$.

A data set was simulated then fitted (using nlme) in Splus 8.2. The process was repeated 1000 times to obtain distributions and bias of model estimates.

Assessments:

1. Distribution of model estimates from bivariate case
2. Bias in model estimates from univariate approach using Y_2 data, compared to bias from the bivariate approach
3. Distribution and bias of model estimates from multivariate case

Bias was calculated as the mean of $100 \cdot (\text{estimate} - \text{true}) / \text{true}$ based on 1000 simulated data sets

RESULTS

Figure 1 shows a typical simulated data set displaying Y_1 (top) and Y_2 (bottom) vs. Time.

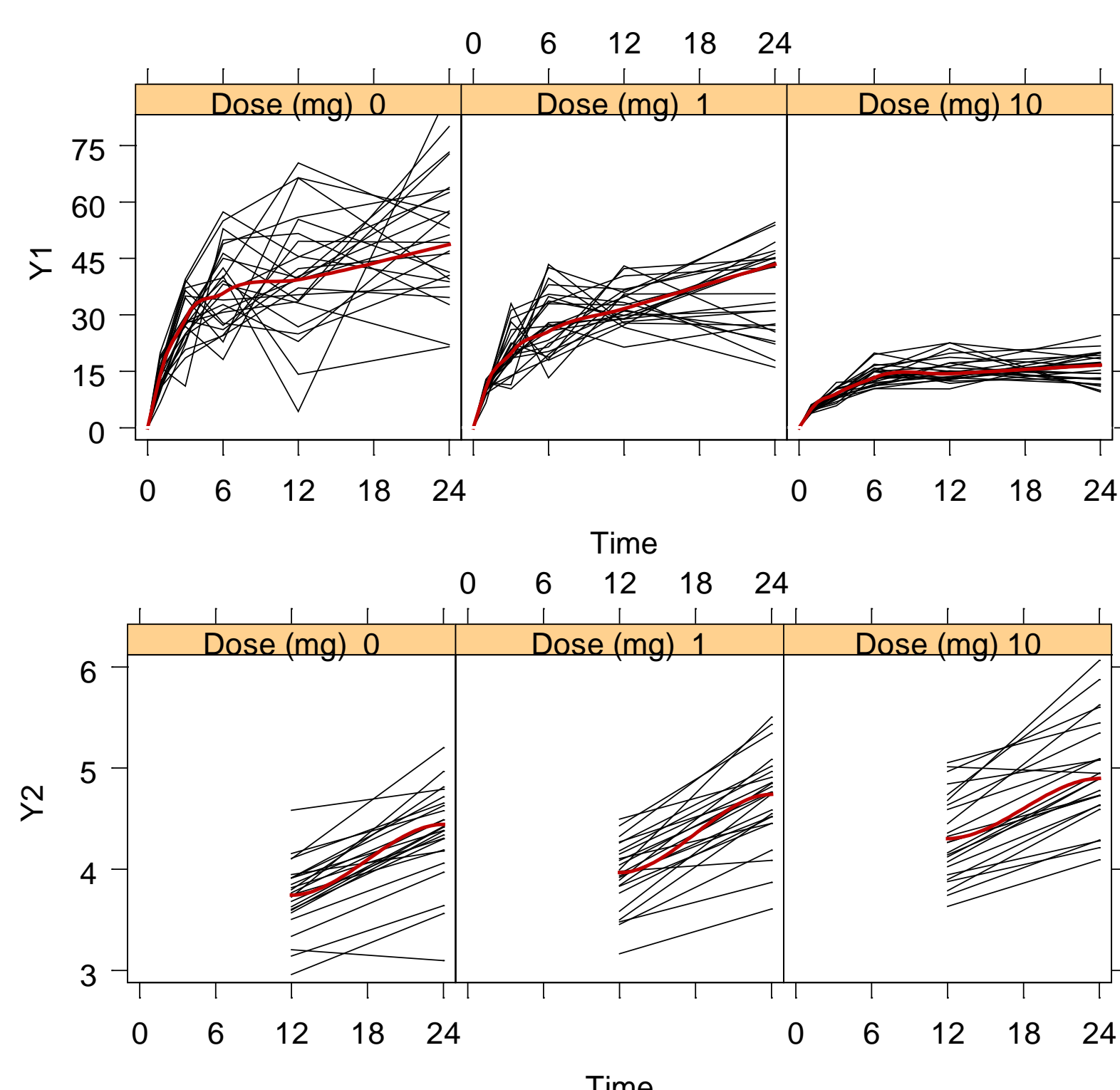


Figure 1: Observed Y_1 and Y_2 (black lines) and loess (red) vs time

RESULTS (Cont.)

The distribution of bivariate model parameters are shown in Figure 2, indicating agreement with true values. Distribution of $Diff. ET_{50} = \log(ET_{50,2}) - \log(ET_{50,1})$ is well centered around true value of 0.5, suggesting ability to test and to accurately estimate differences between shared parameters of endpoints

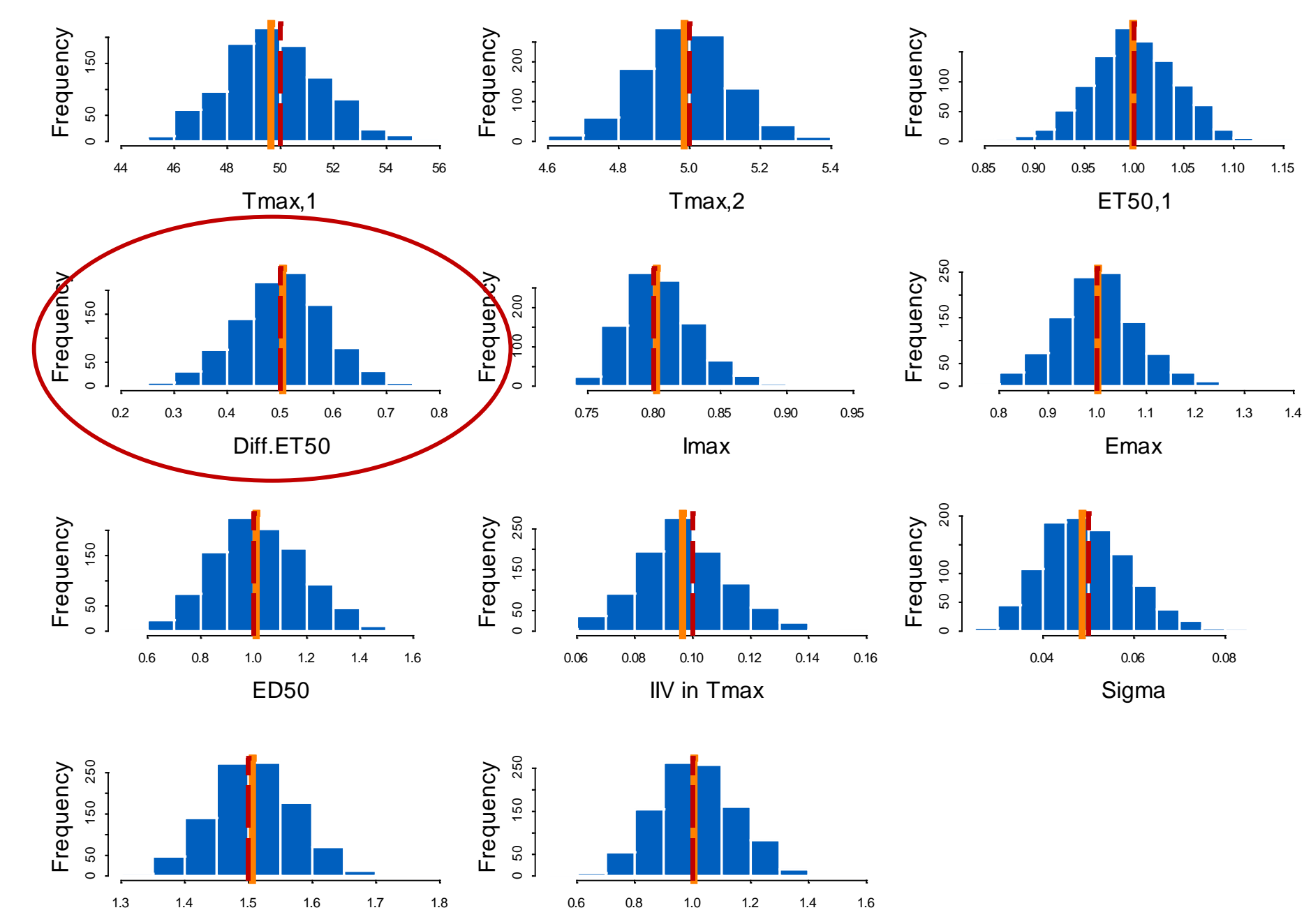


Figure 2: Distribution of bivariate model parameters (1000 runs). Orange vertical line is median of parameter estimate and red line is its true value

Figure 3 (left) shows %Bias in model parameters using the univariate approach for Y_2 against sample size (N). N needs to increase significantly in order to decrease bias to the levels of bivariate model based on N=24, indicated by individual points (lower left corner). Figure 3 (right) shows %Bias in model parameters using the univariate approach for Y_2 against various scenarios of increasing number of measurement time points, as follows: (12 & 24), (6, 12 & 24), (3, 6, 12 & 24), (1, 3, 6, 12 & 24), and (1, 3, 6, 12 & 24 & 48). Again, data points need to increase significantly in order to decrease bias to the levels of bivariate model based of 5 points for Y_1 but only 2 points for Y_2 (indicated by individual points).

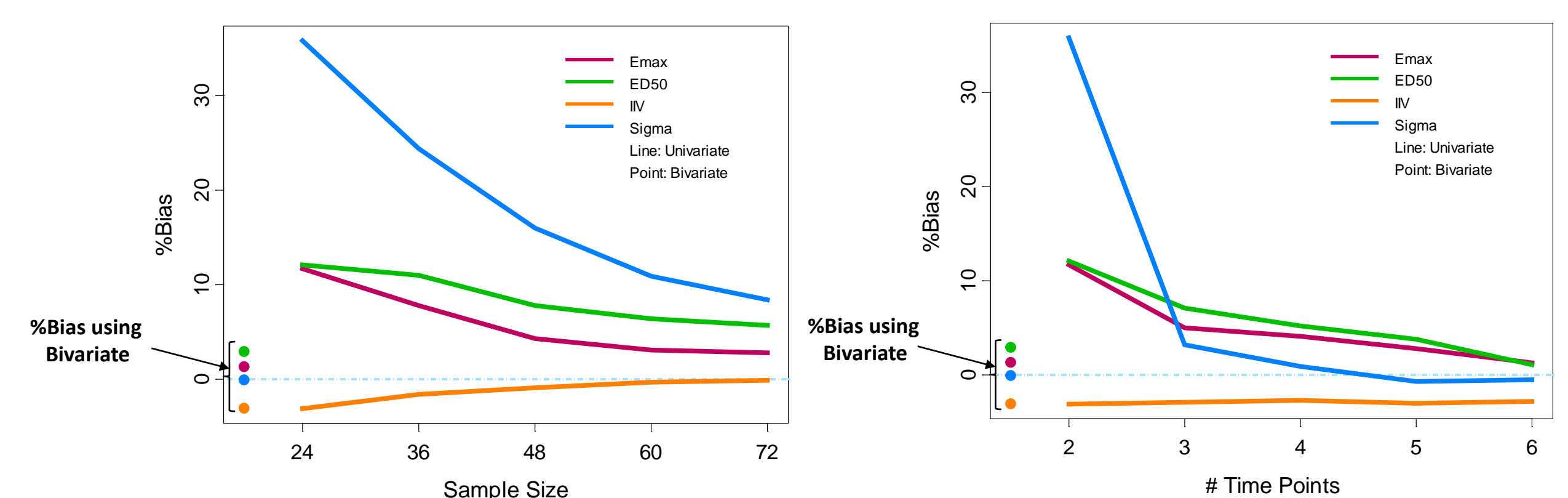


Figure 3: Mean %Bias in model parameters using the univariate approach for Y_2 against sample size (left) and against number of time points (right). Dots represent %Bias in bivariate case

For the multivariate case of 5 endpoints, Figure 4 shows all structural parameters are in agreement with true values (bias less than 1%), suggesting adequacy of the multivariate approach to modeling multiple endpoints

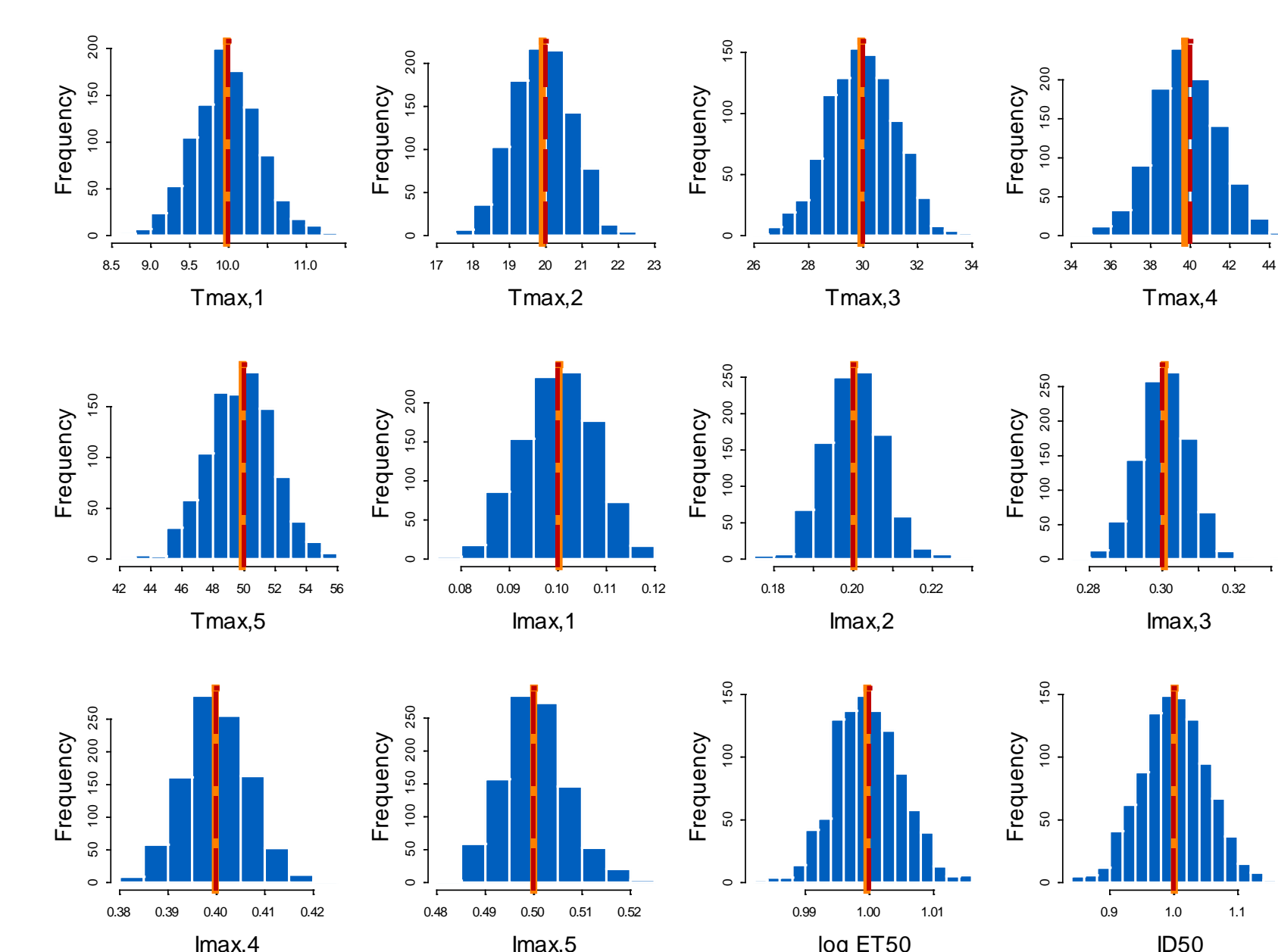


Figure 4: Distribution of multivariate model parameters (1000 runs). Orange vertical line is median of parameter estimate and red line is its true value

DISCUSSION

- The multivariate approach is a general modeling framework applicable to endpoints with different structural and error models. However, model fitting requires software capable of handling heteroscedasticity. The scope of this investigation was limited to the cases discussed here. The robustness of the method will thus depend on amount of information and variability of individual endpoints.
- Modeling multiple endpoints in a single model fit offers the advantage of directly testing magnitude or differences of disease or drug effect on multiple endpoints. For example testing if ET_{50} of disease progression is similar for two biomarkers
- In the cases where endpoints share similar characteristics, but one endpoint is sparsely sampled (due to cost or other practical constraints), then inference on that endpoint produces considerably less bias with the multivariate as compared with the univariate approach
- The value of the multivariate approach is greatest when endpoints share certain features, e.g. same model structure, disease and/or drug effects, with readily available tests between endpoints
- Using the multivariate approach in phase I and early phase II studies, should help optimize, streamline and potentially reduce data collection costs

CONCLUSION

Simultaneously modeling multiple endpoint in a single model fit is a useful and practical approach to investigating drug effects on multiple endpoints.