On the Similarity of Two Dose-Response Curves

Bo Jin * and Kerry B. Barker
Early Oncology Development and Clinical Research, Pfizer Inc.

Abstract
The 2014 FDA draft guidance of clinical pharmacology data to support a demonstration of biosimilarity to a reference product considers clinical pharmacology studies as a critical part in the clinical evaluations on similarity between biosimilars and reference products. A particular issue which remains to be unsettled is on the methodology to demonstrate similarity of two dose response curves. Assuming that dose response curve follows an E-max model, we focus on the discussions on the sensitive dose to detect differences and provide a statistical procedure to test the similarity of two dose response curves.

Key Words: Biosimilarity ; E-max model ; Equivalence

1. Introduction

The 2014 FDA draft guidance of clinical pharmacology data to support a demonstration of biosimilarity to a reference Product [1], describes clinical pharmacology studies as a critical part in the clinical evaluations on similarity between biosimilars and reference products. If clinical pharmacology similarity between products is demonstrated, in some instances this may complete the clinical evaluation, and in others it may support a more targeted clinical development program. As described in the draft guidance, the pathway to demonstrate PK similarity seems to be clear as the bioequivalence approach on bioavailability can be applied, which is similar to that for generic small molecular compounds. However, there remains to be unsettled problems to evaluate the similarity in terms of pharmacodynamics (PD) activities. Among others, the evaluation of dose-response profile plays an important role in the clinical evaluations on similarity, while the methodology to demonstrate the similarity on dose-response profile seems to be unclear.

In this article, we describe an approach to evaluate the similarity of the dose-response profile in a PD marker assuming that the dose-response curve follows an E-max model. Many drug targets are either receptors or enzymes and the corresponding pharmacodynamics activities can often be described by E-max models. In a meta-analysis report of clinical dose response in a large drug development portfolio [2], the E-max model with Hill parameter close to 1.0, describes data very well for most of the compounds in the portfolio.

Denote the reference product as Group 1, and the potential biosimilar as Group 2. let sample size for each group be n, and dose range be $[d_{min}, d_{max}]$ for both groups. Suppose that response of subject j in Group i at Dose d, denoted as as y_{ij}, follows a normal distribution, i.e.,

$$y_{ij} = E_{0,i} + \frac{E_{max,i}d}{d_{50,i} + d} + \varepsilon_{ij}, \tag{1}$$

where $i = 1, 2$, $j = 1, 2, ...n$, $d_{min} \leq d \leq d_{max}$, and $\varepsilon_{ij} \sim i.i.d.N(0, \sigma^2)$. The mean response at Dose d for treatment i is

$$\mu_{i,d} = E_{0,i} + \frac{E_{max,i}d}{d_{50,i} + d}. \tag{2}$$

*Correspondence to: Bo Jin, Early Oncology Development and Clinical Research, Pfizer Inc., 10 Fawcett St., Cambridge, MA 02138. Email: bo.jin@pfizer.com

And then the two treatment difference at Dose d is

$$\Delta(d) = \mu_{1,d} - \mu_{2,d} = E_{0,1} - E_{0,2} + \frac{E_{max,1}d}{d_{50,1}+d} - \frac{E_{max,2}d}{d_{50,2}+d}. \quad (3)$$

Similar to the conventional setup of equivalence testing at a single dose level, we define that two E-max response curves are equivalence if

$$|\Delta(d) = \mu_{1,d} - \mu_{2,d}| < \delta, \quad (4)$$

for any $d_{min} \leq d \leq d_{max}$, where $\delta > 0$ is a pre-specified equivalence margin. Alternatively speaking, the two curves are equivalence if the maximum of $|\Delta(d) = \mu_{1,d} - \mu_{2,d}|$ is less than δ. This paper will tackle the problem of finding the dose that maximizes $|\Delta(d) = \mu_{1,d} - \mu_{2,d}|$ when $d_{min} \leq d \leq d_{max}$ and the statistical inference on testing of equivalence of two E-max dose-response curves. The paper will proceed as follows. Section 2 studies the dose(s) that maximizes $|\Delta(d) = \mu_{1,d} - \mu_{2,d}|$ when $d_{min} \leq d \leq d_{max}$. Section 3 describes a new approach to test the equivalence of two E-max curves. Section 4 uses a simulation study to describe statstical operating characteristics of the new approach, compared to other possible analysis strategies. Section 5 provides some discussions with respect to dose selection and equivalence testing for the clinical development of biosimilar products.

2. The Doses that Maximize the Differences between E-max Response Curves

In this section we find the dose that maximizes $|\Delta(d) = \mu_{1,d} - \mu_{2,d}|$ when $d_{min} \leq d \leq d_{max}$. It can be shown that

$$\frac{\partial(\Delta(d))}{\partial(d)} = \frac{E_{max,1}d_{50,1}}{(d_{50,1}+d)^2} - \frac{E_{max,2}d_{50,2}}{(d_{50,2}+d)^2}. \quad (5)$$

Considering that $d_{50,1} > 0$ and $d_{50,2} > 0$, an unique real root d^* to 5 exists if and only if

$$E_{max,1}E_{max,2} > 0. \quad (6)$$

And the root is

$$d^* = \frac{\sqrt{\zeta}(\sqrt{\zeta} - \sqrt{\eta})d_{50,1}}{\sqrt{\zeta\eta} - 1}, \quad (7)$$

where $\zeta = \frac{d_{50,2}}{d_{50,1}}$ and $\eta = \frac{E_{max,2}}{E_{max,1}}$. Specifically, if $\eta = \frac{E_{max,2}}{E_{max,1}} = 1$ (i.e., $E_{max,2} = E_{max,1}$), then $d^* = \sqrt{\zeta}d_{50,1}$, and if $\zeta = \frac{d_{50,2}}{d_{50,1}} = 1$ (i.e., $d_{50,2} = d_{50,1}$), then $d^* = -d_{50,1}$ which is not a meaningful dose level. In fact it follows from (7) that the sufficient and necessary condition so that $d_{min} < d^* < d_{max}$ is

$$\frac{d_{min}}{d_{50,1}} < \frac{\sqrt{\zeta}(\sqrt{\zeta} - \sqrt{\eta})}{\sqrt{\zeta\eta} - 1} < \frac{d_{max}}{d_{50,1}}. \quad (8)$$

Specifically, if $d_{min} \to 0$ and $d_{max} \to \infty$ then (8) becomes

$$\frac{\sqrt{\zeta}(\sqrt{\zeta} - \sqrt{\eta})}{\sqrt{\zeta\eta} - 1} > 0.$$

It follows that

$$0 < d^* < +\infty \Leftrightarrow \ \zeta > max(\eta, \frac{1}{\eta}) \ or \ \zeta < min(\eta, \frac{1}{\eta}). \quad (9)$$

Theorem 1 *The maximum of* $|\Delta(d) = \mu_{1,d} - \mu_{2,d}|$ *for* $d_{min} \leq d \leq d_{max}$ *occurs only at three points* $d = d_{min}$, d_{max}, *and* d^*, *where* $d_{min} < d^* < d_{max}$ *exits if and only if*

$$\frac{d_{min}}{d_{50,1}} < \frac{\sqrt{\zeta}(\sqrt{\zeta} - \sqrt{\eta})}{\sqrt{\zeta\eta} - 1} < \frac{d_{max}}{d_{50,1}},$$

where $\zeta = \frac{d_{50,2}}{d_{50,1}}$ *and* $\eta = \frac{E_{max,2}}{E_{max,1}}$.

Proof. By Fermat's theorem (stationary points) [3] that the global extrema of a function on a domain occur only at boundaries, non-differentiable points, and stationary points. Theorem 1 then follows considering the condition in (8).

Example 1 *We use an example to describe the above results. Assume that* $E_{0,1} = E_{0,2} = 0$, $E_{max,1} = 1$, $\zeta \in [0.5, 2.0]$, $\eta = 0.8$, $d_{min} = 0.4d_{50,1}$ *and* $d_{max} = 3d_{50,1}$. *By the condition in* (9) *the sufficient and necessary condition for* $0 < d^* < +\infty$ *is* $0.5 \leq \zeta < 0.8$ *or* $1.25 \leq \zeta < 2$. *The dose that maximizes* $|\Delta(d) = \mu_{1,d} - \mu_{2,d}|$ *occurs only at* d_{min}, d_{max} *and* d_*. *Table* 1 *presents the selected results.*

Table 1: An example of the maximum of $|\Delta(d) = \mu_{1,d} - \mu_{2,d}|$ and the corresponding dose: $E_{0,1} = E_{0,2} = 0$, $E_{max,1} = 1$, $\zeta \in [0.5, 2.0]$, $\eta = 0.8$, $d_{min} = 0.4d_{50,1}$ and $d_{max} = 3d_{50,1}$

				$\mu_1 - \mu_2$		
ζ	d^*	$d^* \in (d_{min}, d_{max})$?	at d^*	at d_{min}	at d_{max}	Maximize $\lvert\Delta(d)\rvert$?
0.5	0.360	No	-0.0702	-0.0698	0.0643	d_{min}
0.6	0.302	No	-0.0359	-0.0343	0.0833	d_{max}
0.7	0.192	No	-0.0111	-0.0052	0.1014	d_{max}
0.8	0	No	0	0.0190	0.1184	d_{max}
0.9	-0.340	No	N/A	0.0396	0.1346	d_{max}
1.0	-1	No	N/A	0.0571	0.1500	d_{max}
1.1	-2.615	No	N/A	0.0724	0.1646	d_{max}
1.2	-10.899	No	N/A	0.0857	0.1786	d_{max}
1.3	14.149	No	0.2013	0.0975	0.1919	d_{max}
1.4	5.861	No	0.2085	0.1079	0.2045	d_{max}
1.5	4.239	No	0.2182	0.1173	0.2167	d_{max}
1.6	3.567	No	0.2288	0.1257	0.2283	d_{max}
1.7	3.212	No	0.2395	0.1333	0.2394	d_{max}
1.8	3	No	0.25	0.1403	0.25	$d_{max}(d^*)$
1.9	2.865	Yes	0.2603	0.1466	0.2602	d^*
2	2.775	Yes	0.2702	0.1524	0.27	d^*

3. Testing the Equivalence between Two E-max Dose-Response Curves

We propose a procedure in this section to test the equivalence hypothesis for two E-max dose-response curves as described by (4) using the results from Section 2. The procedure is essentially to test whether the maximum difference of mean responses are within a pre-specified margin of $(-\delta, \delta)$. Since the true model parameters of $E_{0,i}$, $E_{max,i}$ and $d_{50,i}$, where $i = 1, 2$, are unknown and thus need to be estimated from data, d^* will need to be estimated from data and so will the mean responses at the three typical dose levels.

The test procedure is as following:

- Using data to obtain the estimates of of $E_{0,i}$, $E_{max,i}$ and $d_{50,i}$, where $i = 1, 2$ and the estimates are denoted as $\widehat{E_{0,i}}$, $\widehat{E_{max,i}}$ and $\widehat{d_{50,i}}$ respectively. And determine whether $d_{min} < d^* < d_{max}$ exists or not using (8) by substituting $E_{0,i}$, $E_{max,i}$ and $d_{50,i}$ with $\widehat{E_{0,i}}$, $\widehat{E_{max,i}}$ and $\widehat{d_{50,i}}$ respectively.

- Obtaining the estimate of d^* (denoted as $\widehat{d^*}$), as well as the estimates of the mean response differences at d_{min}, d_{max} and d^* (if $d_{min} < \widehat{d^*} < d_{max}$), and determining the location where the maximum of the absolute value the mean response difference occurs by comparing three estimated mean response differences. The estimates of d^* and the mean response differences at d_{min}, d_{max} and d^* (if $d_{min} < \widehat{d^*} < d_{max}$) are as following:

$$\widehat{d^*} = \frac{\sqrt{\widehat{\zeta}}(\sqrt{\widehat{\zeta}} - \sqrt{\widehat{\eta}})\widehat{d_{50,1}}}{\sqrt{\widehat{\zeta}\widehat{\eta}} - 1}, \tag{10}$$

where $\widehat{\zeta} = \frac{\widehat{d_{50,2}}}{\widehat{d_{50,1}}}$ and $\eta = \frac{\widehat{E_{max,2}}}{\widehat{E_{max,1}}}$;

$$\widehat{\Delta(d^*)} = \widehat{E_{0,1}} - \widehat{E_{0,2}} + \frac{\widehat{E_{max,1}}\widehat{d^*}}{\widehat{d_{50,1}} + \widehat{d^*}} - \frac{\widehat{E_{max,2}}\widehat{d^*}}{\widehat{d_{50,2}} + \widehat{d^*}}; \tag{11}$$

$$\Delta(\widehat{d_{min}}) = \widehat{E_{0,1}} - \widehat{E_{0,2}} + \frac{\widehat{E_{max,1}}d_{min}}{\widehat{d_{50,1}} + d_{min}} - \frac{\widehat{E_{max,2}}d_{min}}{\widehat{d_{50,2}} + d_{min}}; \tag{12}$$

and

$$\Delta(\widehat{d_{max}}) = \widehat{E_{0,1}} - \widehat{E_{0,2}} + \frac{\widehat{E_{max,1}}d_{max}}{\widehat{d_{50,1}} + d_{max}} - \frac{\widehat{E_{max,2}}d_{max}}{\widehat{d_{50,2}} + d_{max}}. \tag{13}$$

Note that if $d_{min} < \widehat{d^*} < d_{max}$ does not exist, we may only consider the mean responses difference estimates at d_{min} and d_{max}.

- Obtaining the variance estimate for the estimated largest mean response difference by the Delta method [4] and constructing the associated $(1 - \alpha) \times 100\%$ confidence interval.

- Concluding the equivalence of two E-max curves if the confidence interval of the difference falls within $(-\delta, \delta)$ at the estimated point where the largest absolute difference occurs.

4. A Simulation Study

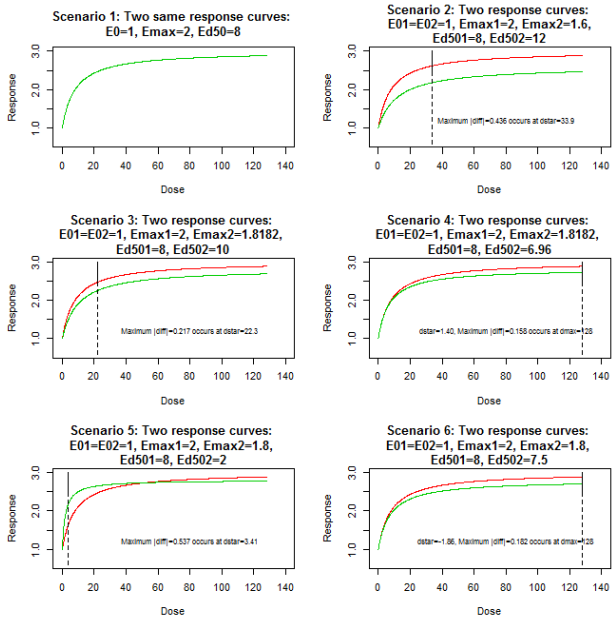
In this section we study the operating characteristics of the proposed testing method. We denote the new testing method as *Emax-new*. The four other methods are: testing equivalence at all the doses using E-max model and we denote this method as *E-max-all*, testing equivalence at only one dose (the highest dose) using E-max model and we denote this method as *Emax-one*, testing equivalence at all the doses using general linear model model (GLM) and we denote this method as *GLM-all*, and testing equivalence at only one dose (the highest dose) using general linear model (GLM) and we denote this method as *GLM-one*.

The equivalence margin is (-0.436, 0.436). The simulated design has eight dose levels: 0, 2, 4, 8, 16, 32, 64 and 128. Six scenarios of the true mean response curves are considered,

Table 2: Six Scenarios of Two E-max Curves: $E_{0,1} = E_{0,2} = 1$, $E_{max,1} = 2$, $d_{50,1} = 8$, $d_{min} = 0$, and $d_{max} = 128$

	Model Parameters		Maximize $\lvert\Delta(d)\rvert$	
Scenario No.	$E_{max,2}$	$d_{50,2}$	at d^* or d_{max}?	Maximum value of $\lvert\Delta(d)\rvert$
1	2	8	N/A	N/A
2	1.6	12	$d^* = 33.9$	0.436
3	1.8182	10	$d^* = 22.3$	0.217
4	1.8182	6.96	$d_{max} = 128$	0.158
5	1.8	2	$d^* = 3.14$	0.536
6	1.8	7.5	$d_{max} = 128$	0.186

Figure 1: Six Simulation Scenarios)



for which the model parameters described in Table 2. The two E-max curves under these scenarios are dispicted in Figure 1.

Note that Scenarios 2 and 5 are corresponding to the case that the two E-max curves are not equivalent, and the error that the equivalence is established in simulation runs is Type I error. Scenario 1 is correponding to the case that the two E-max curves are identical, and the error that equivalence cannot be claimed in simulation runs is Type II error, or alternatively, the probability that we claim the equivalence under this scenario is the power of the test. Other scenarios are corresponding to the cases that the two curve true differences are within the equivalence margin, but they are not identical. We compare the five testing methods in terms of statistical power in these scenarios as well.

Data with the sample sizes of $n = 25$, 50, 75, and 100 per dose each group in the six scenarios following standard normal distribution. The probabilities will be calculated based on 5000 times simulation runs that the 95% confidence intervals for $\Delta(d)$ lies within $(-0.436, 0.436)$ for $0 \leq d \leq 128$, i.e., the probabilities of claiming equivalence. Such probabilities are corresponding to Type I error or power depending on scenario. For the

method of *Emax-all*, the confidence intervals will be constructed at all design dose levels based on E-max model to determine equivalence or not. The similar procedure will be the *GLM-all* method follow which employs GLM instead of E-max model. For both the methods of *Emax-one* and *GLM-one*, only the confidence interval for $\Delta(d_{max})$ will be used to determine equivalence.

Table 3: Probabilities of Claiming Equivalence

		Probabilities of Claiming Equivalence				
Sample Size n	Scenario No.	GLM-one	GLM-all	Emax-one	Emax-all	Emax-new
25	1	0	0	0.1766	0	0.0464
	2	0	0	0.0128	0	0.0006
	3	0	0	0.0940	0	0.0164
	4	0	0	0.1404	0	0.0322
	5	0	0	0.2922	0	0.0006
	6	0	0	0.1250	0	0.0314
50	1	0.1708	0	0.7420	0.2514	0.3582
	2	0.0188	0	0.0376	0.0004	0.0002
	3	0.1102	0	0.3750	0.0074	0.1530
	4	0.1298	0	0.5056	0.1520	0.2492
	5	0.1452	0	0.6964	0.0004	0.0014
	6	0.1156	0	0.4356	0.1326	0.2172
75	1	0.5204	0.0074	0.9308	0.6230	0.6526
	2	0.0296	0	0.0378	0.0010	0.0028
	3	0.2910	0.0002	0.5310	0.2588	0.3410
	4	0.3370	0.0032	0.6826	0.4510	0.4912
	5	0.4478	0	0.8794	0.0006	0.0010
	6	0.2982	0.0012	0.5888	0.4112	0.4360
100	1	0.7308	0.0800	0.9826	0.8340	0.8392
	2	0.0330	0	0.0362	0.0018	0.0056
	3	0.3874	0.0010	0.6572	0.4594	0.5194
	4	0.4904	0.0326	0.8022	0.6760	0.6920
	5	0.6068	0	0.9420	0.0008	0.0010
	6	0.4278	0.0184	0.7392	0.6214	0.6446

The simulation results are presented in Table 3. Under Scenario 2 which is corresponding to the Type I error case (as the true maximum difference is 0.436), the new method of *Emax-new* which tests only three typical dose levels controls Type I error very well, similar to the *Emax-all* and the *GLM-all* methods. Both *Emax-one* and *GLM-one* have the Type I errors greater than the nominal level of 0.025. The Type I error inflation seems not to be very large simply because in this scenario the true $\Delta(d_{max} = 128) = 0.419$, which is not far away from the true maximum difference of 0.436 that occurs at $d^* = 33.9$. The Type I error inflation is seen more clearly in Scenario 5. Since $\Delta(d_{max} = 128) = 0.110$ under Scenario 5, which is away from the the true maximum difference of 0.536 that occurs at $d^* = 3.14$, testing equivalence at only d_{max} as in *Emax-one* and *GLM-one* leads to big Type I error inflation.

As the underlying true models are E-max models, it is not surprising to see the *Emax-*

all method is more powerful than *GLM-all* under these scenarios. The new method of *Emax-new* seems to have higher power compared the method of *Emax-all* when sample size is small ($n = 25$ and $n = 50$). When the sample size is relatively large ($n = 75$ and $n = 100$), the new method of *Emax-new* seems to have the similar power as the method of *Emax-all*.

It is worthy to note that the new method only employs the information of model parameters including their point estimates and the associated variance estimates. In this simulated example, such information is estimated by the data collected at the pre-specified dose levels. But this may not be the only way to get the information of the model parameters. Other source on the information of the model parameters may also be used for the test by this method. For example, if we only have literature data on model parameters without detailed information at each dose level, we can still conduct the equivalence test by the new method. If the response profiles for the reference product and the test product are studied in two separate studies and the dose levels in the two studies are not the same, the equivalence test can be conducted as well. Therefore, the new method has greater flexibility in the application to equivalence test of two Emax response curves, compared to the conventional methods comparing the results on fixed dose levels.

5. Discussion

In this article we have tackled two problems. First problem is which dose level may incur the largest difference between two products assuming that both products have E-max dose-response profiles. We have shown in Section 2 that the maximum difference occurs only at the lowest level (d_{min}), the highest dose level (d_{max}) or the dose level of d^*. We have a sufficient and necessary condition for $d_{min} < d^* < d_{max}$ to exist, and interestingly d^* only depends on the ratio of E_{max} and that of d_{50} between the two E-max models, as seen in Theorem 1.

FDA 2015 guidance on Scientific Considerations in Demonstrating Biosimilarity to a Reference Product [5], suggests that for human pharmacology data assessment, "...when there are established dose-response or systemic exposure-response relationships (response may be PD measures or clinical endpoints), it is important to select, whenever possible, a dose(s) for study on the steep part of the dose-response curve for the proposed product. Studying doses that are on the plateau of the dose-response curve is unlikely to detect clinically meaningful differences between the two products..." The findings in this paper, however, indicate that there may not be no universal answer to the selection of the most sensitive dose to detect the difference between two products. In fact when the dose-response curve is of E-max model, depending on the established dose-response profiles of the two products (i.e.,the E-max model parameters), the most sensitive dose, if 'sensitive' here implies to incur the largest difference, can be the lowest dose, the highest dose and d^* if it exists.

The second problem discussed in the article is how to evaluate the equivalence or similarity between two E-max dose-response curves. Demonstrating the similarity of two dose-response curve is a sensitive test for similarity between products, and "if clinical pharmacology similarity between products is demonstrated, in some instances this may complete the clinical evaluation, and in others it may support a more targeted clinical development program " [1]. In this article, we have proposed a new testing method which is to test the similarity between two products only at the estimated location where the largest difference occurs. This method is flexible in practice with respect to the data source for the equivalence test, and seems to have better statistical operating characteristics when compared to the conventional single-point testing or all-point testing methods in terms of Type I error

control and statistical power in the simulated study.

References

[1] FDA Draft Guidance on Clinical Pharmacology Data to Support a Demonstration of Biosimilarity to a Reference Product, 2014.

[2] Neal Thomas, Kevin Sweeney and Veena Somayaji (2014), Meta-Analysis of Clinical DoseResponse in a Large Drug Development Portfolio, *Statistics in Biopharmaceutical Research*, 6(4), 302-317.

[3] Ron Larson and Bruce H. Edwards (2009), *Calculus* (9th ed.), Brooks/Cole.

[4] Oehlert, G. W. (1992), A note on the delta method, *The American Statistician*, 46(1), 27-29.

[5] FDA Guidance on Scientific Considerations in Demonstrating Biosimilarity to a Reference Product, 2015