Revisiting FDA's 1995 Guidance on Bioequivalence Establishment of Topical Dermatologic Corticosteroids: New Research Based Recommendations

Deniz Ozdin[1,2], Naveen Sharma[3], Jorge Lujan-Zilbermann[4], Philippe Colucci[2], Isadore Kanfer[5,6], Murray P. Ducharme[1,2]

Received, July 14; Accepted, October 30, 2018; Published, November 4, 2018.\

[1] Faculté de pharmacie, Université de Montréal, Pavillon Jean Coutu, 2940 Chemin de la polytechnique, Montréal QC, Canada. [2] Learn and Confirm Inc., 750 Marcel-Laurin, St-Laurent QC, Canada. [3] Cliantha Research Limited. Opposite Pushparaj Towers, Near Judges Bungalows, Bodakdev, Ahmedabad, India. [4] Hill Top Research Limited. 4711 34th Street North, Saint Petersburg, FL-33714, USA. [5] Faculty of Pharmacy, Rhodes University, Grahamstown, South Africa. [6] Leslie Dan Faculty of Pharmacy, University of Toronto, 144 College St, Toronto ON, Canada.

ABSTRACT - Purpose: As per the US FDA guidance issued on June 2, 1995, the establishment of bioequivalence for topical dermatologic corticosteroids is based on comparing the pharmacodynamic (PD) effects of Test and Reference products at the dose duration corresponding to the population *ED50,* determined either by naïve pooled data or nonlinear mixed effect modeling (NLME). The guidance was introduced using a study case example where the expectation maximization (EM) NLME algorithm, as implemented in P-PHARM®, was used. Although EM methods are relatively common, other methods such as the First-Order Conditional Estimation (FOCE) as implemented in the NONMEM® software are even more common. The objective of this study was to investigate the impact of using different parametric population modeling/analysis methods and distribution assumptions on population analysis results. **Methods:** The dose duration-response data from 11 distinct skin blanching blinded pilot studies were fitted using FOCE (NONMEM®) and an EM algorithm (ADAPT5® (MLEM)). Three different *Emax* models were tested for each method. Population PD estimates and associated CV%, and the agreement between model predicted values and observed data were compared between the two methods. The impact of assuming different distributions of PD parameters was also investigated. **Results:** The simple *Emax* model, as proposed in the FDA guidance, appeared to best characterize the data compared to more complex alternatives. The MLEM method in general appeared to provide better results than FOCE; lower population PD estimates with less inter-individual variability, and no variance shrinkage issues. The results also favored ln-normal versus normal distribution assumptions. **Conclusions**: The population *ED50* estimates were influenced by both the type of population modeling methods and the distribution assumptions. We recommend updating the FDA guidance with more specific instructions related to the population approach to be used (EM-like versus FOCE-like methods) and to the normality assumptions that need to be set (ln-normal versus normal distribution).

INTRODUCTION

Bioequivalence (BE) between a Test and Reference (Ref) product is mainly demonstrated using pharmacokinetic (PK) endpoints, such as the area under the curve (AUC) and peak drug concentrations (Cmax), metrics that are related to the rate and extent of exposure of a moiety in the systemic circulation. For locally acting drug products that are not intended to be systemically absorbed such as those administered topically, bioequivalence may be demonstrated using alternative approaches. According to the US FDA 1995 guidance (1), an *in vivo* pharmacodynamic (PD) endpoint is currently an acceptable surrogate to use for BE assessment of topical corticosteroid drug products. The PD response following application of a topical corticosteroid to the skin is its ability to produce a vasoconstriction of the microvasculature of the skin, leading to skin blanching at the site of application. The skin blanching response is then measured visually and/or with a chromameter and is expressed in terms of the Area Under the Effect Curve (AUEC). The use of this PD endpoint for demonstrating BE for topical corticosteroids presumes that skin blanching is sufficiently correlated with the clinical effect, so that two formulations that differ clinically will also differ in terms of skin blanching (2-6).

Corresponding Author: Murray P. Ducharme, Learn and Confirm Inc., 750 Marcel-Laurin, St-Laurent, QC, Canada. Email: murray.ducharme@learnandconfirm.ca

According to the US FDA 1995 guidance, the BE assessment of topical corticosteroids involves the conduct of two separate studies, a pilot and a pivotal study. As a first step, a pilot study is performed solely with the reference listed drug product to establish the response vs. dose-duration relationship, from which PD parameters for use in a pivotal BE study can be determined. To this purpose a topical corticosteroid formulation is applied to the skin of human subjects for differing periods of time, i.e. dose durations. To obtain the PD response for each dose duration, an AUEC is calculated over a time course after drug removal, with time 0 hour representing the time at which the residual drug product was removed until 24 hours later ($AUEC_{(0-24)}$). The calculated PD responses ($AUEC_{(0-24)}$) are plotted as a function of dose duration to obtain the response vs. dose-duration relationship. The relationship is characterized in terms of an *Emax* model. From the relationship, population mean estimates of PD parameters (*Emax* and *ED50*) and other discriminative time points such as D1 and D2 for use in the pivotal bioequivalence study are determined. *Emax* is the maximum PD response or, alternatively, defined as the maximum AUEC. *ED50* is the dose duration required to produce 50% of this maximum PD response. D1 and D2 are dose durations at which approximately 33% and 67% of the maximal effect is produced and are determined as half and double the *ED50*, respectively. Therefore, the calculations of D1 and D2 are directly influenced by the estimated value of the *ED50*.

Within a pivotal study, the BE of a multisource dermatologic corticosteroid is determined by comparing the skin blanching effect produced by the Test formulation to that of the Ref at the population *ED50* estimate identified in the pilot study. This to ensure that the response attributable to this dose duration will fall within a sensitive log-linear region (20% to 80% of *Emax*) of the dose-response curve (7). To make sure that the study is sensitive, only the data from those subjects whose D2/D1 ratios of PD responses meet a specified minimum value of 1.25 may be included in the BE assessment. A robust estimation of the *ED50* from the pilot study is therefore crucial as it may not only affect the Test/Ref ratio of the PD response for BE assessment, but it may also impact on the overall sensitivity of the study in establishing BE.

In order to characterize the response vs. dose-duration relationship and to determine the population *ED50* and *Emax* estimates, the US FDA 1995 guidance recommends to fit the PD response data using either naïve pooled data or a parametric non-linear mixed effect modeling method (NLME). But this latest method requires setting distribution assumptions, which are absent in the guidance. Naïve pooled data does not take into account inter-individual variability when estimating population parameters. As a result, population estimates based on the naïve pooled data method poorly correlate with the observed data, may not accurately represent the study population (8, 9), and should thus not be the method of choice. A NLME is therefore the only option of first choice to be used for the determination of the population parameters of interest such as the *ED50* and the *Emax*. However, there are gaps in the guidance concerning which NLME should or could be used. A case study was included in the guidance when it was issued, and the P-PHARM® software was used by the FDA to fit the skin blanching data to the *Emax* PD model. P-PHARM®, developed at the time by SIMED, was a software package incorporating the Expectation Maximization (EM) method for NLME modeling (10, 11). The EM method, pioneered by Alan Schumitzky and Walker (12), is an estimation method that is based on true likelihood estimation and is incorporated in a large variety of different software such as ADAPT-5® developed and supported by the Biomedical Simulations Resource (BMSR) at the University of Southern California, MONOLIX® by Lixsoft, Phoenix® WinNonlin® by Certara, L.P. (11, 13, 14). Despite this large availability of the EM method for NLME, the NONMEM® software and its FOCE method, originally developed by Beal and Sheiner, is often the method that scientists think of first in terms of NLME and is based on likelihood approximation (15-17). We therefore tested two main methods in this study; the EM algorithm as implemented in the ADAPT-5® software from D'Argenio & Schumitzky (9, 18) and the FOCE algorithm as implemented in the NONMEM® software.

The availability of different population modeling methods is an advantage, as each has its own features and limitations. However, results from different population modeling methods can vary due to their different estimation approaches in data analysis (14, 19, 20). While some studies obtained different population mean estimates when using different population modeling methods (14, 17, 21,

22), some others found comparable results (11, 19, 23).

If employing different population modeling methods results in different population *ED50* estimates, divergent conclusions on BE documentation between topical corticosteroids could be made. Therefore, inconsistency in estimated PD parameters decreases our confidence in BE assessment results for topical corticosteroids. Lack of reproducibility and conflicting results from tape stripping in BE assessment of topical tretinoin gel formulations in two laboratories is the very reason which led to withdrawal of the FDA 1998 Draft Guidance for bioavailability and BE assessment of topical dermatological drug products (24) in May 2002 (25-27). Hence, it is essential for regulatory agencies to update their guidance with instructions recommending more consistent approaches to be followed by pharmaceutical manufacturers.

Similar to other FDA draft guidance documents, the 1995 Corticosteroids guidance is open to various interpretations, and alternative approaches can be used to estimate the PD parameters as long as they comply with the requirements of the regulations. The US FDA Guidance (1) recommends either NLME modeling or naïve pooled data method for skin blanching data analysis to determine the population *ED50 and Emax.* However, it does not specify which type of NLME modeling method should or should not be used, nor does it specify the necessary assumptions for distribution of the PD parameters that need to be set before such an analysis is conducted. Interestingly, in a letter published by FDA in 1998 (28) in response to Demana et al. (29), the absence of any consideration for the nature of the distribution (normal or ln-normal) of population parameters was mentioned among the reasons to discourage the use of naïve pooled data method. However, in the very guidance, no distribution profile was recommended to be assumed for PD parameters when using parametric NLME modeling method. Therefore, one can describe skin blanching data using different PD models, different fitting methods, and different distribution profile assumptions.

As the *ED50* is an essential component in influencing the BE evaluation of topical corticosteroids, its robust estimation is crucial. In this study, we therefore investigated whether different population modeling methods, and different basic fitting assumptions would lead to different population mean estimates of *ED50* and *Emax*. At the time of this research, no other study with real clinical data had been found in the literature investigating the influence of the abovementioned factors on *ED50* estimates for BE assessment of topical corticosteroids. In this study, the objective was to compare population PD estimates obtained from two different NLME population modeling methods as well as different assumptions, and to conclude whether one method/assumption should be prioritized over the other. To ensure practicality and objectivity, we based our analysis on real-life blinded clinical data sets; they were therefore not specifically designed to demonstrate differences among methods. The recommendations that we are putting forward may provide an opportunity for the FDA to update its guidance as well as other regulatory agencies such as Health Canada and European Medicines Agency should they want to consider publishing guidances on the BE assessment of topical corticosteroids using this technique.

METHODS

Data Collected

The data of each study included the PD response for each tested dose duration for one strength of an RLD cream formulation of a corticosteroid. Pharmacodynamic responses were measured in terms of AUEC (unit: scale*time) by means of a chromameter, which was then corrected for baseline and untreated control site for each dose-duration on ventral forearm. A total of 8-10 dose durations were used in each study and the tested dose durations ranged from a minimum of 3 min to a maximum of 360 min. The number of subjects ranged from 16 to 24. The skin blanching data were available from 11 studies. The data were sent blinded in terms of patients and RLD products from Cliantha/Hilltop to Learn and Confirm. It was not possible for the scientists at Learn and Confirm to know what RLD products were tested, but it was mentioned that the same RLD was administered in two of these 11 studies (Studies 6 and 11).

Population Modeling/Fitting of the Response vs. Dose-Duration Data

By use of PD models we attempted to produce the best fit of $AUEC_{(0\text{-}24)}$ versus dose duration data to characterize the response vs. dose-duration relationship from which population mean estimates

of *Emax* (scale*min) and *ED50* (min) could be determined.

For each study the $AUEC_{(0-24)}$ versus dose duration data were fitted twice using two different population modeling methods, the FOCE method as implemented in NONMEM® version VII and the EM method as implemented in ADAPT-5® MLEM version 5.0.53. In each method, the same dataset was fitted separately to three different *Emax* models, the simple *Emax* model as per the FDA guidance, and two modifications to it, an *Emax* model with Hill factor, and an *Emax* model with a minimum Dose-Duration threshold. Goodness of fit measures were used for model discrimination. The best fitting model to the set of observations was selected and was consequently used to compare FOCE and MLEM. The methods were compared in terms of their agreement in population mean estimates of PD parameters and associated inter-individual variability (CV%), and the agreement between model predicted values and observed data. As parametric population modeling methods require certain distribution profile to be assumed, the impact of assuming ln-normal versus normal distribution of PD parameters on analysis results was also investigated.

When using MLEM, each population analysis was run until 1000 population iterations with sampling methods (including important sampling) set at 2000. Results from each MLEM analysis were considered to have started attaining convergence when all PK parameter values had converged graphically (for example, if all PK parameters appeared to have reached stable values starting at population iteration 600). Then the log likelihood estimates for the following 200 population iterations at convergence (in the current example between population iterations 600 and 800) were studied and were verified to not vary between the minimum and maximum estimates by more than 1% for these 200 consecutive iterations. The population iteration number at convergence was then chosen as the first population iteration within that set of 200 that resulted in the exact median convergence estimate (for example if the median value was 1290.20, then if the first population iteration that reached this exact value was 700, then convergence was set to be achieved at population iteration 700). Once the population iteration at convergence was determined, the MLEM analysis was re-run until this exact population iteration in order to get the population parameter values and their associated individual estimates ("*post-hocs*").

When using FOCE, the analyses were permitted to converge automatically by NONMEM® and results were only used if optimization was concluded successfully with a minimum of 3 significant digits.

For both MLEM and FOCE, the values for the initial estimates ("priors") per study for *Emax* and *ED50* were set as the average of the last three AUEC values of all subjects and as the first tested dose duration, respectively.

Structures of the PK-PD Models

As mentioned earlier, the AUEC represents the estimate of the extent of PD response. Therefore, the effect (E) was represented as AUEC in the PK-PD models.

Pharmacological effects being reductions in skin color, the baseline responses were higher than the responses seen following drug applications. Therefore, the AUEC versus dose duration curves have negative slopes, which are reflected by the minus sign in the *Emax* model equations. For all models and for all studies, AUECs are in fact baseline-adjusted control site-corrected $AUEC_{(0-24)}$, *Emax* is the maximum fitted value of AUEC*, ED50* is the dose duration required to produce 50% of the fitted *Emax* value, and *Time* represents each tested dose duration.

Three different models were evaluated for their ability at best characterizing the observed AUEC data from all 11 studies. The simple *Emax* model, as recommended in the FDA guidance, was the base model to which two other *Emax* models were compared. The first one incorporated one additional parameter, a minimum effective dose duration threshold below which no effect could be seen. The second one was the commonly used "sigmoidal *Emax*" model and thereby also incorporated only one additional parameter versus the base model, the "Hill" coefficient. The formulas describing the three models are presented below:

Base model (Simple *Emax* model)

$$AUEC_{(baseline\text{-}adjusted)} = 0 - \frac{\boldsymbol{Emax} \cdot \boldsymbol{Time}}{\boldsymbol{ED_{50}} + \boldsymbol{Time}} \quad (1)$$

Base model with a Minimum Effective Dose Duration Threshold ("MIN")

IF (TIME.LT.MIN) AND IF ((TIME-MIN).LT.0)
THEN Z= 0
ELSE Z=1

$$AUEC_{(baseline\text{-}adjusted)} = 0 - \left(\frac{Emax\,.\,(Time-MIN)}{(ED_{50}-MIN)+(Time-MIN)}\right).Z \qquad (2)$$

Sigmoidal *Emax* model

$$AUEC_{(baseline\text{-}adjusted)} = 0 - \frac{Emax\,.\,Time^{HILL}}{ED_{50}^{HILL}+Time^{HILL}} \qquad (3)$$

For each model, parameters fitted were associated with an inter-individual variability, and the residual variability was fitted using both a proportional and an additive error component. Due to convergence issues, the model with the minimum effective dose duration threshold could only be fitted in NONMEM® with an additive component error.

Statistical Analyses

This study investigated whether the PD parameters were better assumed as ln-normally or normally distributed. In the absence of a known distribution profile, non-parametric statistical tests at a 5% level of significance were used to compare the results (30). The Kruskal-Wallis ANOVA and the Wilcoxon Signed-rank tests were used to compare more than two independent groups and two paired groups, respectively (31, 32). To compare the frequencies of occurrence of a nominal variable (e.g. shrinkage) between two categories (in here, FOCE and MLEM methods), two-sided Fisher's exact test was used (33). SPSS® version 25 (IBM Corp. 2017) and PSI-Plot® version 8.8 (Poly Software International, Inc. 2012, Pearl River, NY) were used for statistical tests and additional generation of graphs.

Model discrimination

According to the law of parsimony, the simplest model should be used preferentially over more complicated models if models fit the data similarly and the standard model discrimination criteria are similar between models (34, 35). The criterion determining the selection of the most appropriate model was the lowest observed value of the Akaike Information Criterion Test (9, 35) when using ADAPT®5 and the lowest observed value of the Minimum Value of the Objective Function (MOF) according to a chi-square distribution ($p<0.05$) (36) when using NONMEM®.

Influence of Different Distribution Assumptions

The impact of different distribution assumptions was investigated based on population mean estimates of PD parameters and their associated inter-individual variability in terms of coefficient of variation (CV%), and their distribution profile. This analysis was performed only with MLEM, as the distribution type must be chosen by the modeller prior to analysis. Therefore, each study dataset was presented to MLEM algorithm twice when using simple *Emax* model; once assuming normal and the other time ln-normal distribution of the PD parameters. This test was not conducted with NONMEM® with FOCE algorithm as the population mean estimate is a geometric mean.

Population PD estimates and associated CV% were compared when normal versus ln-normal distribution were assumed.

Distribution profiles of data were assessed both graphically and numerically for both assumptions. Histograms of PD parameters were generated for graphical assessment. As a numerical test for assessing normality, the Shapiro-Wilk test was used.

Comparison of NONMEM® FOCE and ADAPT®5 MLEM Methods

FOCE and MLEM methods were to be compared only on the most appropriate found structural model, and only when assuming that parameters were ln-normally distributed. Population mean estimates and their associated inter-individual variability in terms of coefficient of variation (CV%), as well as the agreement between model predicted values and observed data (the quality of fit) from the two methods were compared.

Residual variability and Goodness-Of-Fit (GOF) plots were used as the measures of the agreement between model predicted values and observed data. The scatterplots of "observed AUEC versus the corresponding *post-hoc* predicted values" (IPRED versus DV plot) and "weighted residuals versus population predicted AUEC values" (WRES versus PRED plot) were chosen as GOF plots.

Ratios of PD estimates obtained from MLEM versus FOCE and their associated 90% confidence intervals (CIs) were calculated for each study. The two methods were deemed to be similar for specific

parameter estimation if their ratios and 90% CIs were completely within the range of 0.80 to 1.25 interval.

RESULTS

Model Discrimination

The more elaborate models did not improve goodness of fit any further. Therefore, the simple *Emax* model was selected as the most appropriate model to characterize the data, out of the three different models tested. Results of the discrimination process are presented in Table 1.

Influence of Different Distribution Assumptions

Different distribution assumptions caused a statistically significant difference in population mean estimates of *ED50* and *Emax* ($P<0.05$). Medians for population *ED50* estimates were 96.7 and 126.0 min and for population *Emax* estimates were 43.10 and 49.0 scale*min, when normal and ln-normal distribution were assumed, respectively. Inter-individual variabilities around population mean estimates were significantly different when normal versus ln-normal distribution were assumed ($P<0.05$). Population mean estimates and associated CV% under normal and ln-normal distribution assumptions for each study are given in Table 2.

No specific trend could be observed for distribution profile of population *Emax* estimates; in some studies distribution of *Emax* estimates was either normal or ln-normal and in the others neither of them, regardless of initial assumption. For example in study 1, *Emax* estimates were ln-normally distributed while in Study 2 they were neither normally nor ln-normally distributed when ln-normal distribution was assumed.

When ln-normal distribution was assumed, the histogram of population *ED50* estimates displayed a ln-normal distribution. Shapiro-Wilk test results of ln transformed population *ED50* estimates also confirmed the assumed distribution (Figure 1, panel a). When normal distribution was assumed, however, neither normal nor ln-normal distribution profile characterized the obtained *ED50* estimates (Figure 1, panel b).

Table 1. PD model discrimination.

Model	Ref. model	Description	No. of system parameters	MOF	Change in MOF	Selected Model
Base model		"Additive+ proportional" error model	6	1062.6		
Base model + MIN	Base model	"Proportional" error model	7	1063.7	+1.1 (<3.841, P<0.05)	Base model
Sigmoidal *Emax* model	Base model	"Additive+ proportional error model"	8	1070.9	+8.3 (<5.991, P<0.05)	Base model

PD model discrimination in FOCE based on Minimizing the Objective Function (MOF) according to a Chi Square distribution ($P<0.05$)

Model	Ref. model	Description	No. of system parameters	AIC	Selected Model
Base model		"Additive+ proportional" error model	6	1395.8	
Base model + MIN	Base model	"Additive+ proportional" error model	8	1406.3	Base model
Sigmoidal *Emax* model	Base model	"Additive+ proportional" error model	8	1396.9	Base model

PD model discrimination in MLEM based on minimizing the Akaike's Information Criterion (AIC) test

Table 2. Comparison of population mean estimates and associated inter-individual variability (CV%) of each study when different distribution profiles for PD parameters were assumed.

Study	Ln-normal distribution		Normal distribution	
	Geo Mean	**CV%**	**Arith Mean**	**CV%**
1	87.6	128	48.4	49.9
2	255	169	188	57.1
3	501	127	293	43.2
4	38	134	32.6	50
5	37.5	17	44.9	30.7
6	126	120	124	50.7
7	145	166	96.7	55
8	235	91	161	34
9	77	102	50.2	38.2
10	33.7	166	44.2	70.9
11	1220	160	611	49.4

Data represents population geometric (Geo) mean and arithmetic (Arith) mean estimates with associated CV% for *ED50* (min).

Study	Ln-normal distribution		Normal distribution	
	Geo Mean	**CV%**	**Arith Mean**	**CV%**
1	35	19.3	30.2	48.8
2	48.5	36.6	39.2	49.4
3	82.8	26.7	55.1	40.8
4	56	33.3	50.1	37.8
5	16.4	76.1	22.1	62.6
6	58	34.4	51.9	29.7
7	57.1	6.31	43.1	38.7
8	47.8	16.3	42.6	48.3
9	53.3	8.64	43.5	28.9
10	49	31.7	46	32.8
11	30.6	11.7	22.9	51.4

Data represents population geometric (Geo) mean and arithmetic (Arith) mean estimates with associated CV% for *Emax* (scale*min).

As an additional evidence of distribution profile, the geometric mean/median ratios were compared with the arithmetic mean/median ratios of individual PD estimates. When ln-normal distribution was assumed, geometric means and medians of individual *ED50* estimates appeared to be more in agreement with each other than their arithmetic means and medians, implying ln-normal distribution of *ED50* estimates being more likely. In contrast, comparing the ratios for individual *Emax* estimates was unavailing under either assumption (Figure 2).

Comparison of NONMEM® FOCE and ADAPT®5 MLEM Methods

Assuming ln-normal distribution of PD parameters, the base model was used to compare FOCE and MLEM methods, as it was selected as the best fitting model to the set of observations.

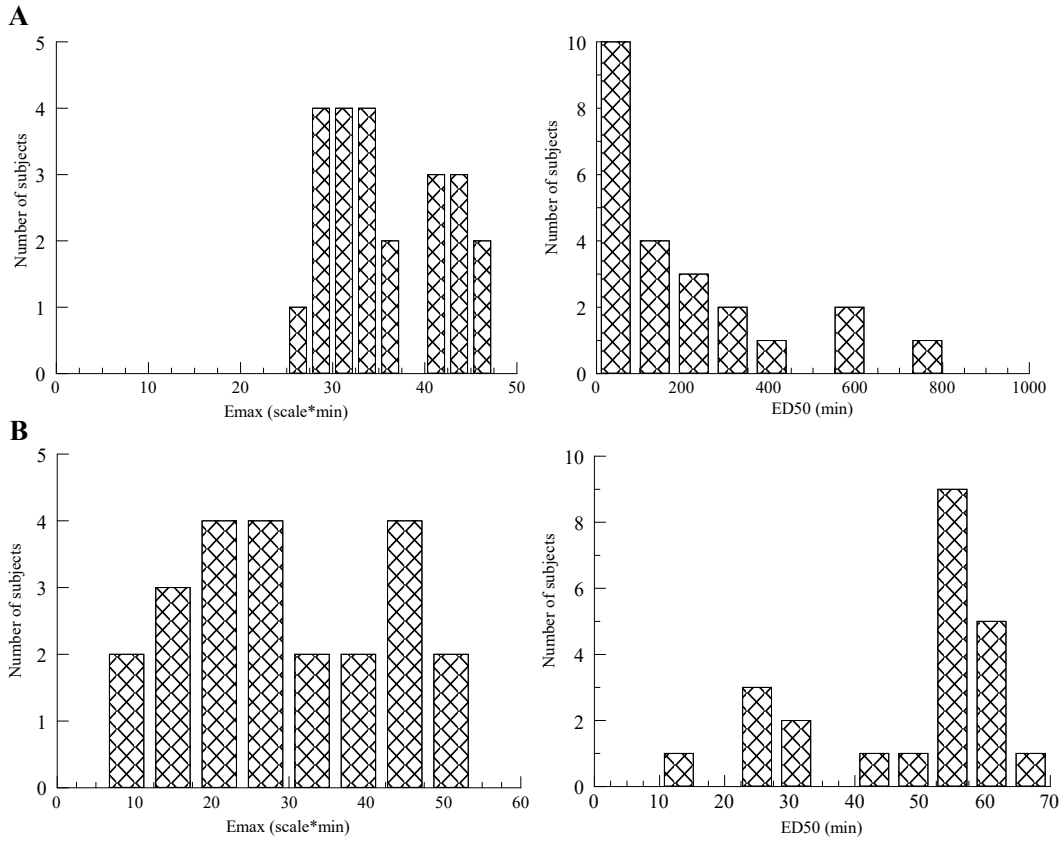

Figure 1. The histogram of population PD estimates for study 1 (n=23). A, when ln-normal distribution was assumed. B, when normal distribution was assumed.

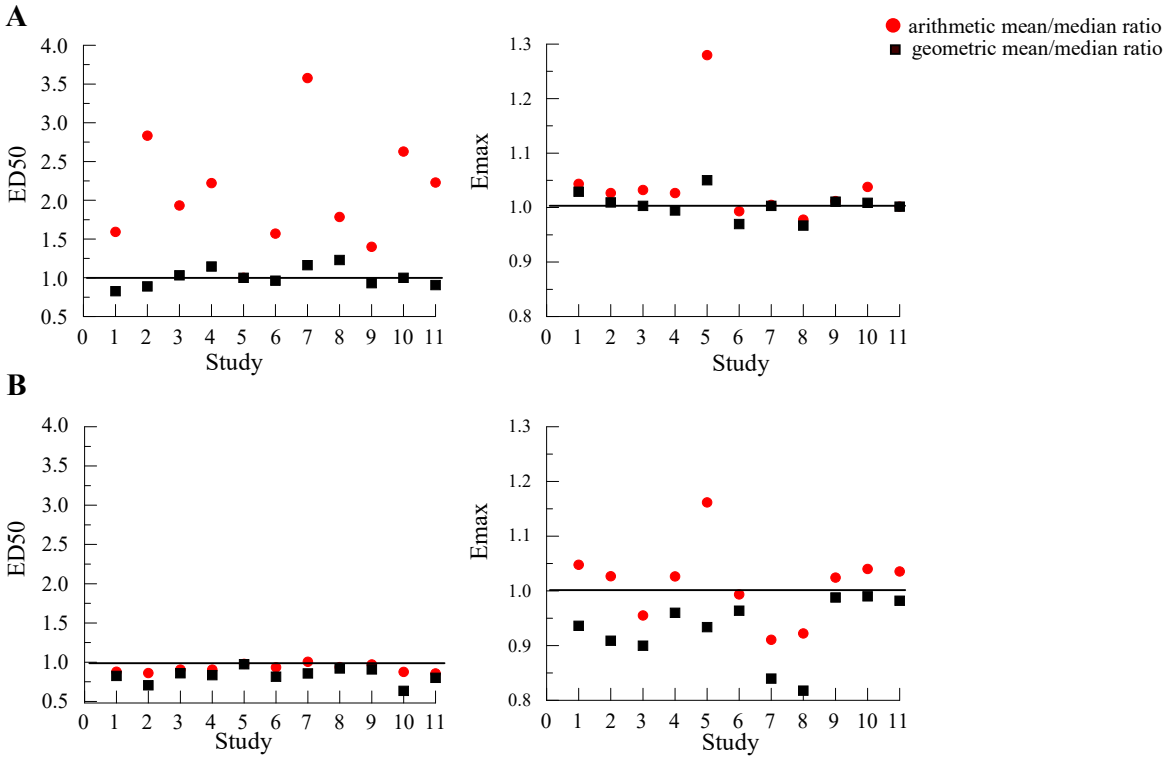

Figure 2. Ratios of arithmetic and geometric mean to the median of *post-hoc* PD estimates. A, when ln-normal distribution was assumed. B, when normal distribution was assumed. The perfect overlay of mean with median is denoted by the long horizontal line.

The NONMEM® FOCE and the ADAPT®5 MLEM methods were found to result in differences. They disagree more often than not between each other, especially in terms of *ED50* and associated CV% rather than *Emax*. Figure 3 shows the dispersion of the ratios of "PD estimates from MLEM versus PD estimates from FOCE" within the range of 0.80-1.25 interval. The 90% CI for the ratios of population *Emax* estimates from all studies, except study 11, fell within the range of 0.80-1.25 which shows the general agreement between two population modeling methods in estimating population *Emax* estimates. However, the dispersion of the *ED50* ratios showed that MLEM estimates in majority tend to be lower than FOCE estimates among which 90% CI for the ratios, in 5 out of 6 studies, completely fell off the lower bound of the range (p=NS, n=11). Population mean estimates and associated CV% for each study are given for FOCE and MLEM in Table 3.

PD parameter variances may be underestimated during the modeling process. When the *post-hoc* estimates are very close to the population values, the variance of *post-hoc* estimates distribution is shrinking towards zero and it becomes difficult to estimate the differences between subjects. This phenomenon is defined as η-shrinkage (sh_η) (36). When the data was fitted with FOCE, 6 out of 22 variance estimates (27%) were associated with shrinkage issue (marked as bolded in Table 3), while there were no instances of this issue when MLEM was used. The number of shrinkage issues with FOCE was significantly higher than with MLEM ($P<0.05$, Fisher's exact test). For the purposes of this study, a severely underestimated variance was associated with an inter-CV<1%.

In study 6 (n=16) and study 11 (n=24) the same RLD formulation was used to characterize the response vs. dose-duration relationship. Therefore, the variability in results of population analysis of repeated applications of the same RLD were determined and compared between MLEM and FOCE. Although MLEM appeared to be more reproducible, the variabilities in population mean estimates were too large with both methods.

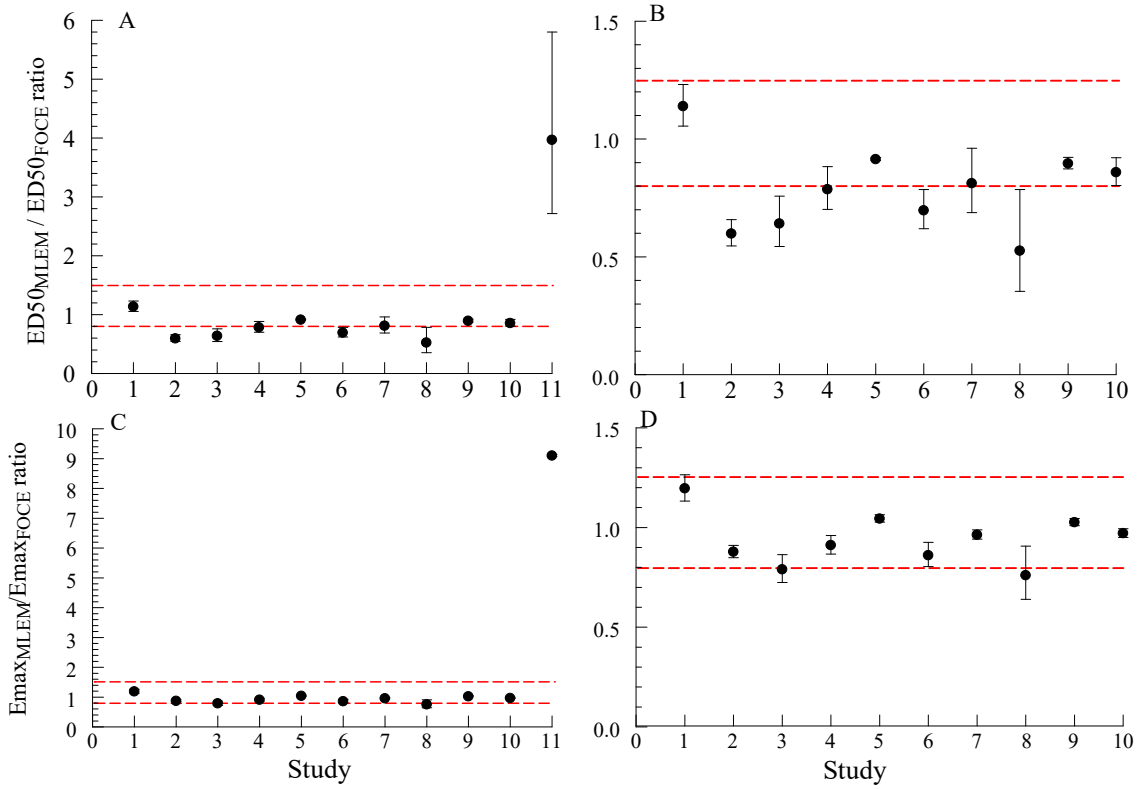

Figure 3. Dispersion of the ratios of "PD estimates from MLEM versus PD estimates from FOCE" within the acceptance range of 0.80-1.25. Each point represents the ratio in each study and black solid bar represent 90% confidence interval for each ratio. Due to the greater variability of PD estimates of study 11, the dispersion profiles were also shown when excluding study 11. A, ratio of *ED50* in all studies. B, ratio of *ED50* in all studies except study 11. C, ratio of *Emax* in all studies. D, ratio of *Emax* in all studies except study 11.

There was no statistically significant difference in residual variability of the fitted model whether FOCE or MLEM were used (P>0.05). When the dispersion of the ratios of residual variability of the model from FOCE versus MLEM was generated, the ratio fell within the range of 0.80-1.25 in 10 out of 11 studies (Figure 4), which implies that the residual variability was similar when different methods were used. In the meantime, GOF plots did not suggest any apparent differences between the two methods. In conclusion, comparative results of residual variability and GOF plots between FOCE and MLEM were indicative of the same overall quality of fit for both methods.

DISCUSSION

Although the BE assessment of many topical formulations still relies on establishing similar clinical efficacy in a comparative clinical study using clinical endpoints in patients, topical corticosteroids are the only exception for which BE can be assessed solely based on an *in vivo* PD study endpoint (i.e., skin blanching) in healthy volunteers. The response vs. dose-duration relationship for topical corticosteroids is demonstrated by application of one strength of drug formulation for varying durations of time as this method has the least manipulation of experimental parameters (37).

Table 3. Comparison of population mean estimates and associated inter-individual variability (CV%) obtained from two NLME modeling methods for each study.

Study	Base model		Base model + MIN		Sigmoidal *Emax* model	
	Geo Mean	CV%	Geo Mean	CV%	Geo Mean	CV%
1	87.6	128	57.6	173	37	72
2	255	169	252.9	169	1950	22
3	501	127	192.1	121	122	63.1
4	38	134	34.3	150	43.4	148
5	37	17.6	62.6	83.4	35.4	16.7
6	126	120	52.6	53.2	71	90.5
7	145	166	124	177	102	156
8	235	91	195	77.7	95.3	51.5
9	77	102	74.5	116	53.6	80.7
10	34	166	31.9	177	663	5.91
11	1220	160	15.6	11.3	221	117

Data represents population Geometric (Geo) mean estimates and CV% of *ED50* (min) with MLEM.

Study	Base model		Base model + MIN		Sigmoidal *Emax* model	
	Geo Mean	CV%	Geo Mean	CV%	Geo Mean	CV%
1	118	320	113	473	170	353
2	219	318	219	320	2850	103
3	373	103	346	63.0	153	95.6
4	35.4	137	34.3	174	37.5	143
5	34.9	**0.55**	35.3	**0.55**	15	751
6	102	130	99.6	132	65.8	101
7	140	247	118	439	146	247
8	132	**0.55**	393	866	10.8	2603
9	78.3	146	71	206	1.4	2.27E+14
10	35.2	288	32.8	373	59.2	463
11	8550	120	2150	250	524	74.3

Data represents population Geometric (Geo) mean estimates and CV% of *ED50* (min) with FOCE. The bolded values represent shrinkage.

The US FDA 1995 Guidance recommends characterizing this PD response in terms of an *Emax* model to determine population *ED50* estimate. Correct estimation of *ED50* is of particular importance as its value will affect validity of the consecutive pivotal study by directly influencing the shorter dose duration calibrator (D1= 0.5**ED50*) and the longer dose duration calibrator (D2= 2**ED50*).

One of the biggest challenges associated with dose duration-response quantification for topical corticosteroids is the determination of the *ED50* of the Ref product as it may not always be reproducible and of adequate reliability (37-40). Due to the importance of accurately estimating population *ED50* in BE assessment of topical corticosteroids, this project aimed to investigate whether different population modeling methods and different basic fitting assumptions would lead to different population PD estimates, and whether a recommendation should be put forward for regulatory agencies to consider for updating and improving guidance documents on the BE assessment of topical corticosteroids.

Given the availability of different types of NLME modeling methods, each study analyst could choose a different method for characterization of AUEC versus dose duration for skin blanching data. It may be acceptable for scientists to use different population modeling techniques if they result in the same estimates. We have, however, seen that this is not the case, and two of the most widely used techniques (e.g., NONMEM® FOCE, and ADAPT5® MLEM) result in significantly different *ED50* estimates. Discrepancies may be indicative of some insufficiency in one or another method or some difficulty arising from a particular dataset, potency, the model, poor starting values, or other sources which require further investigation (17, 19, 21, 41, 42). Some difference in population mean estimates from MLEM and FOCE would be expected due to their different estimation approaches in data analysis, but the differences observed in this study were rather large as the ratios of the *ED50* estimates were rarely within +/-20% of each other.

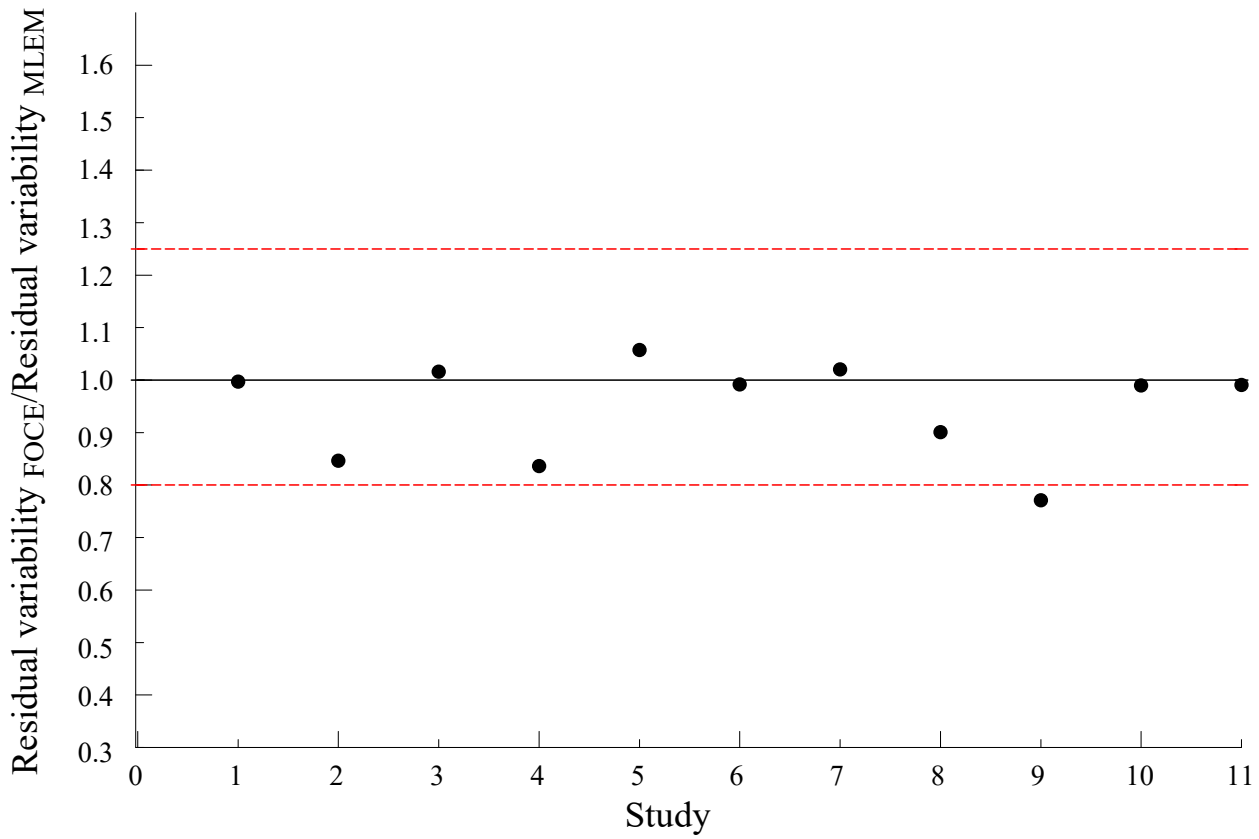

Figure 4. Dispersion of the ratios of residual variability of the model within the range of 0.80-1.25 when the FOCE versus MLEM method was used. Each point represents the residual variability ratio in each study and black solid bar represents equality of the residual variability in the FOCE and MLEM methods.

The impact of population modeling evaluation on *ED50* estimates was previously acknowledged by the US FDA (43).

FOCE and similar estimation methods (FOCE-like methods) implement the estimation of maximum likelihood (ML) to solve the nonlinear problems. Therefore, the population mean estimates from FOCE-like methods are based on model approximation and not true ML estimators (20, 44). In addition, FOCE-like methods first fit the data by obtaining population mean estimates followed by a conditional second step with individual data estimates (*post-hocs*) in an iterative fashion. The fixed effects and random effects are fitted simultaneously with respect to population mean and variability estimates as well as the residual variability (11, 45). MLEM and similar estimation methods (EM-like methods), on the other hand, compute maximum likelihood with an iterative approach that involves 2 repetitive steps; an expectation step (E-step) and a maximization step (M-step). Since the linear approximation is replaced by importance sampling-based estimation method, the parameters obtained are true/exact ML estimates (11, 20, 44, 46, 47). In the E-step, parameter variables are estimated using the latest predicted parameter values and the observed data (Bayesian estimation of the individual parameters). In the M-step, parameter values are estimated and updated to maximize the log-likelihood function in the E-step (estimation of population parameters). These new values are then reused for the subsequent iteration (46, 47). In the MLEM algorithm, ML is combined with an EM algorithm (10, 11, 48-50). EM-like methods, similar to FOCE-like methods, fit the data in an iterative fashion, but in a different order; they first compute the individual estimates (*post-hocs*) followed by the population estimates.

In this study, population PD analyses of topical RLD corticosteroid formulations were conducted for 11 different PD effect study data, using two different population modeling methods and different distribution profile assumptions. A satisfactory model was developed in both methods based on goodness of fit measures (MOF and AIC) and residual variability. In both methods, the simple *Emax* model described the skin blanching data better than two modified forms of it which were tested in this study. In general MLEM appeared to provide "better" results than FOCE did. MLEM provided lower population PD estimates (Figure 3) with less variability, and no issue of variance shrinkage (Table 3).

The simple *Emax* model (hyperbolic model) as suggested by the FDA 1995 guidance remains at this time preferable to more complex/modified models. This is in agreement with previously published results. Demana et al. investigated the suitability of two different PD *Emax* models to describe skin blanching data as the result of topical corticosteroid application; simple *Emax* model and sigmoidal *Emax* model. They concluded the chromameter data were best described by the simple *Emax* model (29). Later on, other studies also found that sigmoidal *Emax* model did not improve the model fit (51, 52).

Both FOCE and MLEM algorithms are designed to estimate the central tendency of population data using parametric methods (17). Therefore a certain type of distribution assumption is required to be made prior to performing the population analysis. Owing to the different shapes of them, the central tendency values of normal and ln-normal distributions vary; while in normal distribution the arithmetic mean and median overlay, in ln-normal distribution geometric mean and median overlay (53, 54). As a consequence, different distribution profiles may cause difference in central tendency values, in our case, population mean estimates of *Emax* and *ED50*. The FDA 1995 Guidance does not recommend a particular type of distribution (normal versus ln-normal) to be assumed for running NLME population modeling. Based on our literature search, only two studies mentioned the type of distribution profile that they assumed within population analysis for fitting AUEC versus dose duration data (51, 52). Both studies assumed normal and ln-normal distribution for *Emax* and *ED50* parameters, respectively. However, neither mentioned the basis for this choice. In this study, the distribution profile of data was assessed both graphically and numerically. The initial distribution assumptions significantly affected the population *Emax* and *ED50* estimates and associated CV%, as well as the distribution profile of population *ED50* estimates. When ln-normal distribution was initially assumed, population *ED50* estimates appeared to be ln-normally distributed. When normal distribution was assumed, however, neither normal nor ln-normal distribution profile could be achieved. Investigating the influence of the two assumptions on distribution profile of population *Emax* estimates was unpersuasive. Given the fact that *ED50* serves as an essential component in influencing BE evaluation of

topical corticosteroids, the results of analysis for *ED50* estimates were prioritized in making a conclusion. The results, therefore, suggested that assuming ln-normal distribution of the PD parameters should be favored over normal distribution for skin blanching studies.

Although not written in the guidance, lower *ED50* estimates should be preferred to ensure that the comparison between a Test and Ref in pivotal study remains in the sensitive portion of the dose duration response curve. For this reason, the population modeling method which provided lower estimates was considered to be preferable in this study. Numerical comparison of the results of each study suggested that the estimates were different between the two methods. The dispersion of the ratios of PD estimates from MLEM versus PD estimates from FOCE (Figure 3) shows that *ED50* estimates from MLEM in majority appeared to be lower than FOCE estimates. This is in agreement with previously published results. Staatz and Tett also found population mean estimates lower when they used EM-like versus FOCE-like method to fit the blood concentration–time data of orally administered tarcrolimus (17). Many other studies also reported different results from FOCE-like and EM-like methods (14, 21, 22), but none had used skin blanching data of topical corticosteroids.

The estimated inter-individual CVs around each PD parameter by different methods for the same dataset may be indicative of the uncertainty associated with the population estimate of the PD parameter (55). Given the example provided in the 1995 US FDA guidance, no data indicating the inter-individual variability associated with the population *ED50* estimate appears to be required by the agency. As a consequence one could never know the inter-individual variability around the estimated *ED50* and, therefore, on the D1 and D2 time points. In this study, we assumed that a better population modeling method would be associated with detectable (e.g. No shrinking issues) but lesser inter-individual variability around the *ED50* estimate. Results showed that inter-individual variabilities around population mean estimates generally appeared lower in MLEM than in FOCE and no variance shrinkage was observed with MLEM while it was an issue with FOCE. This finding was in agreement with the results previously found by Colucci et al (42). However, the CV% associated with *ED50* estimates were significantly higher than CV% associated with *Emax* estimates for both methods (median of 137.34% and 128.0% versus 25.27% and 26.70% when FOCE and MLEM were used, respectively). Given that we often "hear" that sponsors experience difficulties at determining the *ED50* values in pilot skin blanching studies, the greater variability in *ED50* estimates seen was not surprising. Tsai et al also found a higher CV% for population *ED50* estimates than for population *Emax* parameters (≈40-90% versus ≈10-27%) in a dose duration-response characterization study for clobetasol 17-propionate (52).

In this study, the same RLD formulation was used in study 6 and 11. Therefore population mean estimates obtained from the two studies were expected to be comparable. In contrast, a large difference was observed between the population mean estimates for both NLME methods, implying neither FOCE nor MLEM were of adequate reproducibility for this dataset. In two surveys conducted by FDA, one of which was on 88 ANDAs with vasoconstrictor BE studies submitted from January 1992 to April 2015 (56), high variability and lack of reproducibility and consistency in PD response/skin blanching data were reported alike as the difficulties with topical corticosteroids experienced by sponsors (39, 43, 51).

Wide variation in residual variabilities of ≈30% to 250% was observed in our results with both methods; median residual variabilities with the *Emax* model were 61.89 and 67.20 when FOCE and MLEM were used, respectively, implying high uncertainty left after the data were fitted and therefore overall low degree of model fit to the *Emax* model. The literature search showed that the data of skin blanching study is inherently associated with high degree of variability. High variability in AUEC data have been found by many investigators (28, 30, 38, 57, 58). In a study performed by Smith et al, there were extensively large standard deviations about the mean values of AUEC at each time point with no differentiation between the means. Approximately 20-50% variability were found in skin blanching data (30). As such, an intra-individual variability of 60%-139% was found in a study performed by Singh et al (51) which was inversely related to dose duration and to the potency of the dermatologic corticosteroid product. In a more recent study conducted by Lehman and Franz (58), variability ranging from 78-126% were reported in the skin blanching data which were fitted to the *Emax* model. The authors, however, did not specify which fitting method was used. The variability in the mass of formulation

applied to the skin, different application sites along the forearm, chromameter probe manipulations, ambient temperature, relative humidity, posture, one application site for each dose duration, and adjusting the chromameter readings for the baseline are some sources contributing to the high variability in skin blanching data (29, 51, 59-61).

In conclusion, this study demonstrated that using different population modeling methods and different assumptions regarding distribution of PD parameters affected the population estimation of the PD parameters and their associated variability. Regardless of the population modeling method used, more complex versions of the simple *Emax* model did not appear to be necessary to describe skin blanching data better. The study results suggested that EM-like methods may provide better population *ED50* estimates of skin blanching data. It also suggested that ln-normal distribution should be assumed for the distribution of the *ED50* parameter.

As population *ED50* estimates play a critical role in the BE assessment of topical corticosteroid products, any difference in estimated PD parameters could influence the outcome of BE evaluation for these products. Due to the availability of several methods for performing population modeling and their parametric approach in data analysis, updating the US FDA 1995 Guidance with more specific instructions related to the population approach and normality assumptions, would favor a more consistent approach to be followed by pharmaceutical manufacturers, and would increase the confidence in BE assessment results of these products.

ACKNOWLEDGMENTS

The studies described in this manuscript were sponsored by Learn and Confirm Inc., Cliantha Research Limited and Hill Top Research Limited.

REFERENCES

1. US FDA Guidance for Industry: Topical Dermatologic Corticosteroids: In Vivo Bioequivalence, 02/June/1995. Last accessed 05/Aug/2016, at: http://www.fda.gov/ohrms/dockets/dockets/04p0206/04p-0206-ref0001-08-FDA-Guidance-for-Industry-06-1995-vol3.pdf.
2. Hauck WW. Bioequivalence studies of topical preparations: statistical considerations. International journal of dermatology. 1992;31 Suppl 1:29-33.
3. Stoughton RB. The vasoconstrictor assay in bioequivalence testing: practical concerns and recent developments. International journal of dermatology. 1992;31 Suppl 1:26-8.
4. Cornell RC, Stoughton RB. Correlation of the vasoconstriction assay and clinical activity in psoriasis. Archives of dermatology. 1985;121(1):63-7.
5. Gibson JR, Kirsch JM, Darley CR, Harvey SG, Burke CA, Hanson ME. An assessment of the relationship between vasoconstrictor assay findings, clinical efficacy and skin thinning effects of a variety of undiluted and diluted corticosteroid preparations. The British journal of dermatology. 1984;111 Suppl 27:204-12.
6. Pershing LK, Silver BS, Krueger GG, Shah VP, Skelley JP. Feasibility of measuring the bioavailability of topical betamethasone dipropionate in commercial formulations using drug content in skin and a skin blanching bioassay. Pharm Res. 1992;9(1):45-51.
7. Samara E, Granneman R. Role of population pharmacokinetics in drug development. A pharmaceutical industry perspective. Clin Pharmacokinet. 1997;32(4):294-312.
8. Bonate PL. Recommended reading in population pharmacokinetic pharmacodynamics. The AAPS Journal. 2005;7(2):E363-E73.
9. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: Routine clinical pharmacokinetic data. Journal of Pharmacokinetics and Biopharmaceutics. 1983;11(3):303-19.
10. Aarons L. Software for Population Pharmacokinetics and Pharmacodynamics. Clinical Pharmacokinetics. 1999;36(4):255-64.
11. Mentre F, Gomeni R. A two-step iterative algorithm for estimation in nonlinear mixed-effect models with an evaluation in population pharmacokinetics. Journal of biopharmaceutical statistics. 1995;5(2):141-58.
12. Walker S. An EM algorithm for nonlinear random effects models. Biometrics 1996; 52(3):934-944.
13. Ng CM. Novel hybrid GPU-CPU implementation of parallelized Monte Carlo parametric expectation maximization estimation method for population pharmacokinetic data analysis. Aaps j. 2013;15(4):1212-21.
14. Plan EL, Maloney A, Mentre F, Karlsson MO, Bertrand J. Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods for dose-response models. Aaps j. 2012;14(3):420-32.
15. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical

pharmacokinetic data. J Pharmacokinet Biopharm. 1980;8(6):553-71.
16. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. J Pharmacokinet Biopharm. 1981;9(5):635-51.
17. Staatz CE, Tett SE. Comparison of two population pharmacokinetic programs, NONMEM and P-PHARM, for tacrolimus. Eur J Clin Pharmacol. 2002;58(9):597-605.
18. Schumitzky A. EM Algorithms and two stage methods in pharmacokinetic population analysis. In: Advanced methods of pharmacokinetic and pharmacodynamic system analyses. Vol. II. D'Argenio, DZ editor. Plenum Press; New York:1995. p 145-160.
19. Roe DJ. Comparison of population pharmacokinetic modeling methods using simulated data: results from the Population Modeling Workgroup. Statistics in medicine. 1997;16(11):1241-57; discussion 57-62.
20. Bauer RJ, Guzy S, Ng C. A survey of population analysis methods and software for complex pharmacokinetic and pharmacodynamic models with examples. Aaps j. 2007;9(1):E60-83.
21. Bennett JE, Wakefield JC. A comparison of a Bayesian population method with two methods as implemented in commercially available software. J Pharmacokinet Biopharm. 1996;24(4):403-32.
22. Girard P, Mentré F. A comparison of estimation methods in nonlinear mixed effects models using a blind analysis. In: Population Approach Group in Europe P, editor. Population Approach Group in Europe, PAGE; Pamplona, Spain. Pamplona, Spain: Population Approach Group in Europe, PAGE; 2005.
23. McLachlan AJ, Tett SE. Pharmacokinetics of fluconazole in people with HIV infection: a population analysis. British journal of clinical pharmacology. 1996;41(4):291-8.
24. US-FDA Guidance for Industry: Topical Dermatologic Drug Product NDAs and ANDAs-In vivo Bioavailability, Bioequivalence, In Vitro Release, and Associated Studies, 1998. Bethesda (MD): Center for Drug Evaluation and Research, FDA..
25. D.P. Conner, Differences in DPK Methods, http://www.fda.gov/ohrms/dockets/ac/01/slides/3804s2_01_conner/index.htm, Advisory Committee for Pharmaceutical Sciences Meeting, Center for Drug Evaluation and Research (CDER), Food and Drug Administration (FDA), Rockville, MD, November 29, 2001.
26. Franz, TJ. Study #1, Avita Gel 0.025% vs Retin-A Gel 0.025%, Advisory Committee for Pharmaceutical Sciences Meeting, Center for Drug Evaluation and Research (CDER), Food and Drug Administration (FDA), Rockville, MD, November 29, 2001.
27. Pershing, LK. Bioequivalence assessment of three 0.025% tretinoin gel products: Dermatopharmacokinetic vs. clinical trial methods, Advisory Committee for Pharmaceutical Sciences Meeting, Center for Drug Evaluation and Research (CDER), Food and Drug Administration (FDA), Rockville, MD, November 29, 2001.
28. Singh GJ, Fleischer N, Lesko L, Williams R. Evaluation of the proposed FDA pilot dose-response methodology for topical corticosteroid bioequivalence testing. Pharm Res. 1998;15(1):4-7.
29. Demana PH, Smith EW, Walker RB, Haigh JM, Kanfer I. Evaluation of the Proposed FDA Pilot Dose-Response Methodology for Topical Corticosteroid Bioequivalence Testing. Pharmaceutical Research. 1997;14(3):303-8.
30. Smith EW, Haigh JM, Walker RB. Analysis of chromameter results obtained from corticosteroid-induced skin blanching. I: Manipulation of data. Pharm Res. 1998;15(2):280-5.
31. Lynn KP. Assessment of Topical Corticosteroid-Induced Skin Blanching Response Using the Visual Mckenzie-Stoughton and Colorimetric Methods. Drug information journal. 1995;29(3):923-34.
32. Holford NH, Sheiner LB. Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models. Clin Pharmacokinet. 1981;6(6):429-53.
33. Fisher RA. On the interpretation of chi square from contingency tables, and the calculation of P. Journal of the Royal Statistical Society. 1922;85(1):87-94.
34. Wade JR, Beal SL, Sambol NC. Interaction between structural, statistical, and covariate models in population pharmacokinetic analysis. J Pharmacokinet Biopharm. 1994;22(2):165-77.
35. Yamaoka K, Nakagawa T, Uno T. Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations. Journal of Pharmacokinetics and Biopharmaceutics. 1978;6(2):165-75.
36. Savic RM, Karlsson MO. Importance of shrinkage in empirical bayes estimates for diagnostics: problems and solutions. AAPS J. 2009;11(3):558-69.
37. Pershing LK, Lambert L, Wright ED, Shah VP, Williams RL. Topical 0.050% betamethasone dipropionate. Pharmacokinetic and pharmacodynamic dose-response studies in humans. Archives of dermatology. 1994;130(6):740-7.
38. Walker RB, Haigh JM, Smith EW. Application of the Minolta Chromameter to the Assessment of Corticosteroid-lnduced Skin Blanching. In: Schwindt DA, Maibach HI, editors. Cutaneous Biometrics. Boston, MA: Springer US; 2000. p. 295-305.
39. Keith G. Bioequivalence of topical corticosteroids: Design and data analysis challenges with the vasoconstrictor assay. 6th World Congress on Bioavailability & Bioequivalence: BA/BE Studies

Summit; August 17-19, 2015; Chicago, USA: Journal of Bioequivalence & Bioavailability; 2015. p. 31.
40. Gibson JR, Kirsch J, Darley CR, Burke CA. An attempt to evaluate the relative clinical potencies of various diluted and undiluted proprietary corticosteroid preparations. The British journal of dermatology. 1983;109 Suppl 25:114-6.
41. Aarons L, Balant LP, Mentre F, Morselli PL, Rowland M, Steimer JL, et al. Population approaches in drug development. Report on an expert meeting to discuss population pharmacokinetic/pharmacodynamic software. Eur J Clin Pharmacol. 1994;46(5):389-91.
42. Colucci P, Grenier J, Yue CS, Turgeon J, Ducharme MP. Performance of Different Population Pharmacokinetic Algorithms. Therapeutic Drug Monitoring. 2011;33(5):583-91.
43. Ren K, Braddy A, Wang R, Caramenico H, Conner D. FDA Perspective on Current Challenges with the Use of Pharmacodynamic Endpoint Evaluation of Bioequivalence of Topical Dermatologic Corticosteroids. AAPS; November 10-13, 2013; San Antonio2013
44. Davidian M, Giltinan DM. Nonlinear models for repeated measures data. New York: Chapman & Hall; 1995.
45. Seng Yue C and Ducharme MP: empirical models, mechanistic models, statistical moments, and noncompartmental analysis. In Shargel L, Yu Andrew BC (eds). Applied Biopharmaceutics and Pharmacokinetics. 7th ed. The McGraw-Hill Inc. 2016. .
46. Gomeni R, Pineau G, Mentre F. Population kinetics and conditional assessment of the optimal dosage regimen using the P-PHARM software package. Anticancer research. 1994;14(6a):2321-6.
47. Kinowski J-M, Bressolle F, Rodier M, Augey V, Fabre D, Richard JL, et al. A Limited Sampling Model with Bayesian Estimation to Determine Inulin Pharmacokinetics Using the Population Data Modelling Program P-PHARM. Clinical Drug Investigation. 1995;9(5):260-9.
48. Wang X, Schumitzky A, D'Argenio DZ. Nonlinear Random Effects Mixture Models: Maximum Likelihood Estimation via the EM Algorithm. Computational statistics & data analysis. 2007;51(12):6614-23.
49. Wang J, Weiss M, D'Argenio DZ. A note on population analysis of dissolution-absorption models using the inverse Gaussian function. Journal of clinical pharmacology. 2008;48(6):719-25.
50. Wang X, Schumitzky A, D'Argenio DZ. Population Pharmacokinetic/Pharmacodyanamic Mixture Models via Maximum a Posteriori Estimation. Computational statistics & data analysis. 2009;53(12):3907-15.
51. Singh GJ, Adams WP, Lesko LJ, Shah VP, Molzon JA, Williams RL, et al. Development of in vivo bioequivalence methodology for dermatologic corticosteroids based on pharmacodynamic modeling. Clinical pharmacology and therapeutics. 1999;66(4):346-57.
52. Tsai J-C, Cheng C-L, Tsai Y-F, Sheu H-M, Chou C-H. Evaluation of in vivo bioequivalence methodology for topical clobetasol 17-propionate based on pharmacodynamic modeling using Chinese skin. Journal of pharmaceutical sciences. 2004;93(1):207-17.
53. Heath DF. NORMAL OR LOG-NORMAL - APPROPRIATE DISTRIBUTIONS. Nature. 1967;213(5081):1159.
54. Weippl G, Pantlits.M. LOG-NORMAL OR NORMAL DISTRIBUTION AS A BASIC REQUIREMENT FOR KNOWLEDGE OF NORMAL VALUES1970.
55. Graaf PHvd. Introduction to Population Pharmacokinetic/Pharmacodynamic Analysis With Nonlinear Mixed Effects Models. CPT: Pharmacometrics & Systems Pharmacology. 2014;3(12):e153.
56. J K, Y W, K R, R W, A C B, N T, et al. Evaluation of Quality of Submission of Vasoconstrictor Assay Bioequivalence Studies for Topical Dermatologic Corticosteroid Products in Abbreviated New Drug Applications (ANDAs). AAPS; October 25–29; Orange County Convention Center, Orlando, Florida.2015.
57. Pershing LK, Bakhtian S, Poncelet CE, Corlett JL, Shah VP. Comparison of skin stripping, in vitro release, and skin blanching response methods to measure dose response and similarity of triamcinolone acetonide cream strengths from two manufactured sources. Journal of pharmaceutical sciences. 2002;91(5):1312-23.
58. Lehman PA, Franz TJ. Assessing Topical Bioavailability and Bioequivalence: A Comparison of the In vitro Permeation Test and the Vasoconstrictor Assay. Pharmaceutical Research. 2014;31(12):3529-37.
59. Kirsch J, Gibson JR, Darley CR, Barth J, Burke CA. Forearm site variation with the corticosteroid vasoconstrictor assay. The British journal of dermatology. 1982;106(4):495.
60. Ale SI, Laugier JP, Maibach HI. Spacial variability of basal skin chromametry on the ventral forearm of healthy volunteers. Archives of dermatological research. 1996;288(12):774-7.
61. Sommer A, Veraart J, Neumann M, Kessels A. Evaluation of the vasoconstrictive effects of topical steroids by laser-Doppler-perfusion-imaging. Acta dermato-venereologica. 1998;78(1):15-18.