Pharmaceutical Research, Vol. 16, No. 7, 1999

Letter to the Editor

Truncated Area Under the Curve as a Measure of Relative Extent of Bioavailability: Evaluation Using Experimental Data and Monte Carlo Simulations

To The Editor:

A recent paper by Gaudreault *et al.* in the October 1998 issue of *Pharmaceutical Research (1)* presents some new data on the use of truncated areas under the curve to determine equivalence of extent of drug absorption for drugs with long half-lives. The authors evaluated experimental data from 123 bioequivalence studies conducted on long half-life drugs and compared how the use of various truncated areas affected the determination of bioequivalence compared to use of the conventional areas under the curve to time infinity (AUCinf).

The truncated areas that were compared to AUCinf were: (a) area under the curve to 2 times Tmax (AUC2*Tmax); (b) area under the curve to 3 times Tmax (AUC3*Tmax); (c) area under the curve to 4 times Tmax (AUC4*Tmax); and (d) area to last quantifiable concentration (AUClqc). As one might expect, the degree of concordance between AUCinf and the other metrics increased with the length of the sampling period. However, since there was a 1.6% difference between even AUClqc and AUCinf, the authors concluded that there is no perfect agreement among AUC metrics, even when drug concentrations are measured for relatively long periods of time after dosing.

The authors also presented simulated studies from which the data were used to calculate the relative probabilities (i.e. power) of meeting the bioequivalence criteria (90% confidence limits of 80–125%) for extent of absorption under various experimental conditions when using different metrics of area [AUCinf, AUClqc, area under the curve to 72 hours (AUC(0–72)), and area under the curve to 8 hours (AUC(0–8)]. Cases investigated for the one-compartment model were:

1. *Low LOQ*-baseline conditions with limit of quantitation set equal to $1/25^{th}$ of predicted mean Cmax of reference formulation
2. *High LOQ*-baseline conditions with limit of quantitation set equal to $1/10^{th}$ of predicted mean Cmax of reference formulation
3. *Highly variable Model I*--12.7% intra subject coefficient of variation (CV) on ke (elimination rate constant).
4. *Highly variable Model II*--29.7% intra subject CV on F*D/V (where F = bioavaialibility, D = dose, V = volume)
5. *Highly variable Model III*--26.3% intra subject CV on ke,

and, for the two-compartment model:

1. Model IV--disposition half-life of 10 hr. and terminal half-life of 198 hr.
2. Model V--disposition half-life of 4 hr and terminal half-life of 88 hr.

Both two-compartment models were investigated at LOQ values of 1 and 10 μg/ml. For Model IV, this allowed drug quantitation to 3 and 1 terminal half-lives, respectively, and for Model V, to 5 and 1 terminal half-lives, respectively.

The results for the one-compartment model scenarios exhibited no consistent pattern related to power, although AUCinf and AUClqc had the highest power in 3 of the 5 cases studied. For the two-compartment model scenarios, however, the hierarchy for power in each case was AUC(0–8) > AUC(0–72) > AUClqc > AUCinf. The article concludes with some very convincing arguments for the use of truncated areas, especially AUC(0–72), which has been proposed by the Health Protection Branch of Canada in a draft guidance as a metric for determining equivalent extent of drug absorption for drugs with long half-lives.

Results from the simulations clearly indicate that the performance (i.e. power) of an AUC metric is related to the kinetics of the individual drug being tested. This circumstance is potentially troublesome for regulatory agencies, especially if a generic product is found to be bioinequivalent based upon the favored AUC(0–72) metric. Under these conditions, it is likely that the sponsor would "shop" for that alternative metric believed most likely to provide an advantageous result, thereby requiring the regulatory agency to determine which metric would be most appropriate for that particular drug product and to enforce that determination.

An article in *Pharmaceutical Research* by Kharidia, *et al.* (2) presents an alternative approach to the use of truncated areas to measure extent of absorption for long half-life drugs. Data from simulated bioequivalence studies were used to categorize drugs into two groups. In the first group, no matter which last time point (24, 48, 72, or 96 hours post-dose?) was chosen to calculate the truncated AUC, the resultant 90% CI's were in accordance. A second group of drugs did not exhibit this characteristic. The simulations were based upon the work of Endrenyi and Tothfalusi (also cited by Gaudreault *et al.*) who proposed using AUC(0–24) as an alternative metric for long half-life drugs.

Kharidia *et al.* (2) found that the most important factors predicting concordance of the various truncated AUC confidence intervals were (a) the relative similarity of intrasubject variation for 90% CI's independent of the last drug concentration time used in the calculations and (b) low inter-subject variability in the times of last quantifiable drug concentration. These circumstances resulted in similar amounts of data being lost from each subject's terminal concentration-time profile, resulting in concordant 90% CI's among the various truncated AUC's.

Conversely, these researchers found that in those drugs that did not exhibit concordance, the level of intrasubject variation, as indicated by the 90% CI's, increased with the length of time the drug was measured post-dosing. This was a result of high and low clearance subjects within the population, which

0724-8741/99/0700-1144$16.00/0

caused the subjects to exhibit their last measurable plasma concentration at many different time points.

Amiodarone and danazol were used as models for drugs exhibiting concordance and non-concordance, respectively, between truncated AUC values and AUCinf.

Gaudreault *et al.*, (1) and Endrenyi and Tothfalusi, (3) defined bias as [(Truncated AUC-AUCinf) /AUCinf] and the authors used it to measure the accuracy of their simulations. The bias was approximately 3% for AUC(0–24), indicative of good accuracy for the metric, using the one-compartment model with a test-to-reference ratio of 3 for the absorption rate constant (Kat/Kar). Simulations presented in the in-press article (2) duplicating and expanding upon the studies of Endrenyi and Tothfalusi, yielded similar results for power. However, a different and potentially more relevant estimate of accuracy was employed by Kharidia *et al.* (2). Accuracy was assessed by determining the number of times the simulated true mean ratio for fraction available (Fa Test /Fa Reference) was within the calculated 90% CI. The data showed quite clearly that even though AUC(0–24) had higher power than AUC(0–96), the AUC(0–96) metric was generally more accurate.

The work of Kharidia *et al.* (2), then, provides the following scientific and regulatory advantages over the approach favored by Gaudreault, *et al.*

1. The use of truncated AUC values is restricted to those drugs having little inter/intra-subject variation in time of last quantifiable plasma/serum concentration. For this group of drugs, no matter which truncated AUC metric was chosen, the bioequivalence outcome would not be changed from that obtained using the current metrics of AUClqc and AUCinf. Although the metrics are consistent with our current measures of AUClqc and AUCinf, truncation would be preferred since it requires fewer subject samples to be analyzed.
2. The current metrics of AUClqc and AUCinf would still be used for all other drug types.

The application of this approach would place no additional burdens upon sponsors of generic drug products or the drug regulatory agencies and would support present and past determinations of bioequivalence.

Andre J. Jackson
and
Larry A. Ouderkirk
Office of Pharmaceutical Science,
CDER, US FDA,
Rockville, Maryland

REFERENCES

1. J. Gaudreault, D. Potvin, J. Lavigne, and R. L. Lalonde. Truncated area under the curve as a measure of relative extent of bioavailability: Evaluation using experimental data and Monte Carlo simulations. *Pharm. Res.* **15**:1621–1629 (1998).
2. J. Kharidia, A. J. Jackson, and L. A. Ouderkirk. Use of truncated AUC to measure extent of drug absorption in bioequivalence studies: Effects of drug absorption rate and elimination rate variability on this metric. *Pharm. Res.* **16**:130–134 (1999).
3. L. Endrenyi and L. Tothfalusi. Truncated AUC evaluates effectively the bioequivalence of drugs with long half-lives. *Int. J. Clin. Pharmacol. Ther.* **35**:142–150 (1997).

THE AUTHORS REPLY:

We wish to address some of the comments of Drs. Jackson and Ouderkirk concerning our recent publication (1) as well as their own publication (2) in this journal.

They report the results of two bioequivalence studies with danazol and amiodarone which are very consistent with the 123 studies that we reported in our publication. Table 1 in their publication shows the Test/Reference ratio of AUCs truncated at different times and the root mean square error from the analysis of variance (a measure of the residual "intra-subject" variability). These values change very little over time in the post-absorption phase which is precisely what we observed for most of the 123 studies that we reported (see Figure 1 of our publication). Similar results were also reported by Midha *et al.* for 10 different studies and 24 different analytes (3,4).

Drs. Jackson and Ouderkirk correctly point out that in our simulations the relative power (i.e., probability of demonstrating bioequivalence) for AUCs truncated at different times varied with the different simulation scenarios. This is to be expected and it was one of the reasons for evaluating the different simulation scenarios. They further indicate that sponsors may therefore "shop" for alternative AUC metrics if the results of a particular study fail to meet the bioequivalence criterion based on the proposed AUC_{0-72}. We agree with them that such "shopping" would be clearly inappropriate. However, there is a solution to this potential problem. The regulatory agency simply has to clearly articulate what AUC metric must be used to assess relative extent of bioavailability (e.g., the Canadian Health Protection Branch proposal to use AUC_{0-72} for long half-life drugs). This is analogous to the current situation with AUC_{inf} being used by the FDA and sponsors are not allowed to arbitrarily select another AUC metric when they fail with AUC_{inf}. Furthermore, the blood sampling scheme (e.g., collecting samples only up to 72 hours after the dose) and the parameters used for the determination of bioequivalence should be specified in the study protocol, thereby precluding any "shopping" after the results become available. Therefore, we don't see why there should be any more regulatory difficulty with the use of a specific truncated AUC than with the current use of AUC_{inf}.

Jackson and Ouderkirk agree with our proposal concerning the use of truncated AUCs but wish to restrict it to drugs that exhibit little inter/intra-subject variability in the time at which the lower limit of quantitation of the assay is reached. The basis for this restriction apparently comes from their study results with danazol and simulations based on the danazol data (2). The authors argue that there is some fluctuation in the confidence intervals reported for danazol AUCs truncated at different times after the dose. However, the fluctuations in the width of the confidence intervals are relatively small and are expected for studies with relatively few subjects, such as bioequivalence studies. The observed fluctuations with danazol are consistent with the fluctuations that we observed in the 123 studies included in our report. Furthermore, it should be noted that their results for danazol include some unusual values because the Test/Reference ratio for AUC_{0-36} is outside its own 90% confidence interval (it is unclear if the reported ratios are based on geometric least-squares means). It is important to recall that the number of subjects (actually the number of subjects with a measured AUC parameter in both study periods) is also an important determinant of the width of the confidence

interval for the AUC ratio. Everything else being equal, if there are many missing values for AUC_{inf} because the terminal half-life cannot be estimated for certain subjects, the confidence interval for the ratio of AUC_{inf} will be wider as a consequence of the fewer degrees of freedom and larger t statistic used to calculate the confidence interval. Under such circumstances, AUC_{inf} is a less appropriate metric since the bioequivalence conclusions are based only on a subset of the study population. This may explain the apparently different Test/Reference ratio and wider confidence interval for danazol AUC_{inf} reported in their Table 1 because some of the other truncated AUC metrics have actually greater residual "intra-subject" variability and yet produce more narrow confidence intervals (e.g., AUC_{0-t}, AUC_{0-96} and AUC_{0-72}).

Their simulation set B was used to support the recommendation to restrict the use of truncated AUCs. However, it should be noted that this simulation is for a drug with a 4-fold higher rate of absorption (Ka) and 25% higher extent of absorption for the Test compared to the Reference formulation, and a mean half-life of 24 hours. The mean half-life would actually be about 17 hours in the subpopulation with the 40% higher clearance that was included in simulation set B. Once inter-subject and intra-subject variability are added around the above mean values, there will be many subjects with relatively short half-lives, as evidenced by the fact that the lower limit of quantitation was reached as early as 24 hours after the dose in some cases. Thus, it is very important to emphasize that this particular simulated drug is very different from the long half-life drug (mean of 100 hours) that was simulated in our study. For drugs with shorter half-lives and a 4-fold difference in Ka (as in simulation set B), there is an expected bias in the ratio of AUCs if they are truncated too soon after the dose.

The authors proposed a measure of bias/accuracy based on the number of times the true AUC ratio is included in the calculated confidence interval for each simulated study. This is a measure of the "coverage" of the calculated confidence interval, as described by other investigators (5). While confidence interval coverage is an interesting measure to report, it is important to note that it depends on both bias (difference between the mean estimate and the true ratio) and standard error of the AUC ratio estimate (the latter is a function of the residual variability from the analysis of variance). Decreased confidence interval coverage below the nominal level of 90%, as reported for their simulation set B, may be due to bias, inappropriately low estimate of residual variability, or both. It would be interesting for the authors to report each factor separately so that the cause of the lower confidence interval coverage can be better evaluated. For example, a biased estimate of the AUC ratio would explain why the power for both AUC metrics in their Table 3 is more than 8-fold higher than the expected 5% (recall that the true Test/Reference ratio of bioavailability in the simulations was 1.25), as was observed in Table 2 with simulation set A. Higher power for either AUC metric in Table 3 should not be considered a positive attribute because we would like no more than 5% of studies to pass the bioequivalence criterion when the true ratio of bioavailability is 1.25 (i.e., the consumer risk is controlled by the two one-sided t tests with alpha = 5%).

Figure 2 of their article clearly demonstrates how the apparent limitations of truncated AUCs are eliminated when the ratio of Ka/Ke is higher (i.e., for drugs with longer half-lives). For example, confidence interval coverage is much closer to the expected value of 90% and the power much closer to the expected value of 5% when the Ka/Ke ratio is above 20. Most importantly, truncated AUCs performed well with the higher ratios of Ka/Ke, despite the increased variability that was included in the simulations. Therefore, the apparent limitations of truncated AUCs with simulation set B were really a consequence of the relatively short half-life and low Ka/Ke ratio used in the simulations.

In summary, the results reported by Jackson and Ouderkirk are generally in agreement with our own results that truncated AUCs are a good measure of relative extent of bioavailability, particularly for drug with long half-lives. Their simulation set B demonstrated a potential limitation of truncated AUCs when there is a large difference in Ka between formulations, however this was a consequence of the relatively short half-life used in the simulations.

Jacques Gaudreault
Genentech
and
Richard L. Lalonde
Parke-Davis Pharmaceutical Research

REFERENCES

1. J. Gaudreault, D. Potvin, J. Lavigne, and R. L. Lalonde. Truncated area under the curve as a measure of relative extent of bioavailability: Evaluation using experimental data and Monte Carlo Simulations. *Pharm. Res.* **15**:1621–1629 (1998).
2. J. Kharidia, A. J. Jackson, and L. A. Ouderkirk. Use of truncated areas to measure extent of absorption in bioequivalence studies: Effects of drug absorption rate and elimination rate variability on this metric. *Pharm. Res.* **16**:130–134 (1999).
3. K. K. Midha, J. W. Hubbard, M. Rawson, and L. Gavalas. The application of partial areas in assessment of rate and extent of absorption in bioequivalence studies of conventional release products: experimental evidence. *Eur. J. Pharm. Sci.* **2**:351–363 (1994).
4. K. K. Midha, J. W. Hubbard, and M. J. Rawson. Retrospective evaluation of relative extent of absorption by use of partial areas under the plasma concentration versus time curves in bioequivalence studies on conventional release products. *Eur. J. Pharm. Sci.* **4**:381–384 (1996).
5. D. B. White, C. A. Walawander, Y. Tung, and T. H. Grasela. An evaluation of point and interval estimates in population pharmacokinetics using NONMEM analysis. *J. Pharmacokinet. Biopharm.* **19**:87–112 (1991).