

Bioequivalence Testing for Highly Variable Drug Products (HVDP)

James R. Johnson, Cary, NC, USA

ABSTRACT

Bioequivalence (BE) studies investigate and compare the pharmacologic characteristics of different formulations of a drug product with respect to the rate and extent of exposure to a new formulation (Test) and a reference listed formulation (RLD). The pharmacokinetic parameters of maximum concentration observed (C_{max}), time to maximum concentration (T_{max}), and area under the concentration curve (AUC) are the most common parameters for assessing rate and extent of drug absorption in BE studies, and provide clues to the impact of outlier concentrations. Highly variable drug products (HVDP) are drugs whose rate and extent of absorption shows large dose-to-dose variability within the same subject. HVDP's are generally defined as those drugs whose intra-patient coefficient of variation (C_{max} and/or AUC) is approximately 30% or greater. In this paper methods associated with determining bioequivalence for a HVDP are summarized from the literature. The regulatory constraints associated with evaluating HVDP are briefly summarized and discussed. The complication associated with selecting an appropriate sample size for bioequivalence testing of HVDP is summarized in the presence of regulatory constraints and design criteria.

INTRODUCTION

Highly variable drugs' have been defined as those drugs for which the within-subject variability (WSV) equals or exceeds 30% of the maximum concentration (C_{max}) and/or the area under the concentration versus time curve (AUC). Despite the fact that highly variable drugs are generally safe with flat dose response curves, the bioequivalence of their formulations is a problem because the high variability means that large numbers of subjects are required to give adequate statistical power. Highly variable drug products are poor quality formulations where high within-formulation variability (e.g. tablet to tablet variability) poses a problem rather than high innate WSV of the drug itself. A further problem caused by high variability is that a subset of the population may respond differently to the two formulations producing a significant subject x formulation interaction.

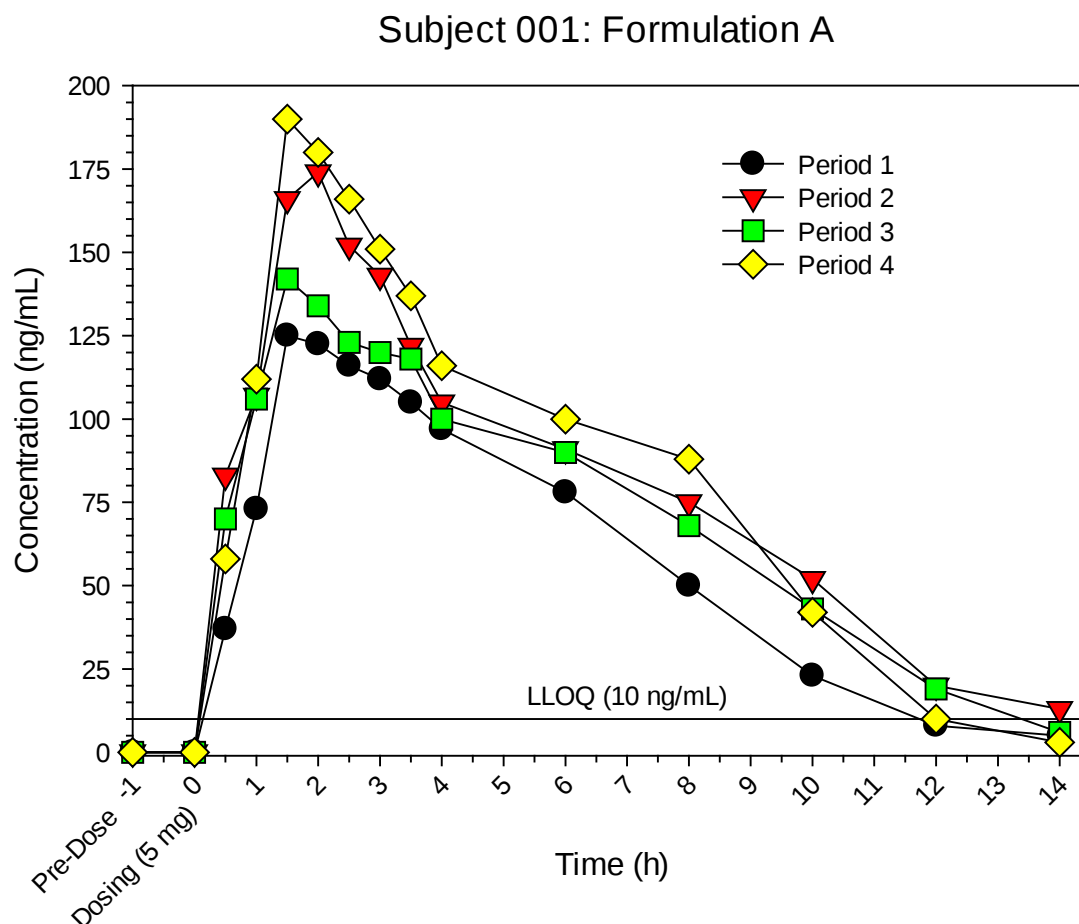
Practical examples are shown using replicate designs. The methods proposed to deal with the problems posed by highly variable drugs include:

- (i) Drug regulatory jurisdictions states that the 90% confidence interval (90% CI) around the test to reference geometric mean ratio (GMR) is required to fit with bioequivalence acceptance limits of 0.8 - 1.25 for both C_{max} and AUC.
 - a. The WSV for single point estimation of C_{max} is often greater than that for AUC. One strategy therefore is not to require a 90% CI for C_{max} of drugs that do not exhibit a toxicity associated with C_{max} and merely require the GMR to fall within the acceptance limits.
- (ii) To arbitrarily broaden the bioequivalence acceptance limits.
 - a. For example, to permit a sponsor to justify the use of wider limits e.g. the 90% CI around the GMR of C_{max} values might be required to fit within acceptance limits of 0.75 - 1.33 or even 0.70 - 1.42.
- (iii) A more systematic approach would be to broaden the acceptance limits by scaling to either the residual variance from a 2-period design or to the WSV of the reference product in a replicate design.
 - a. Subsequent evaluations of scaling procedures have demonstrated that smaller numbers of subjects are required for bioequivalence studies on formulations of highly variable drugs.
 - b. A disadvantage of scaling is that the method is less sensitive to differences between the means compared with un-scaled treatment, such that the GMR may prove to be unacceptably low or high. This possibility has led to a suggestion that the GMR must fall within acceptance limits of 0.8 - 1.25 in scaled treatments.
- (iv) A similar method is to scale the metric rather than the acceptance limits. This method was proposed by the United States' Food and Drug Administration in the context of Individual bioequivalence, but may also be applied to average bioequivalence.
- (v) To carry out bioequivalence studies at steady state whenever a multiple dose regimen is ethically acceptable for healthy volunteers.
 - a. This solution is based on the observation that high variability in a single dose study tends to be dampened at steady state, thus increasing statistical power. Drug regulators have not favored this approach on the grounds that bioequivalence testing should be based on the most discriminating test possible.
- (vi) Finally the use of metabolite data has been proposed since in many (but by no means all) cases, metabolite is less highly variable than that of the parent drug. This subject remains controversial except when the administered substance is a prodrug which is converted by metabolism into the active drug.

INTRA-SUBJECT VARIABILITY

Variability within a sample can be best described through the use of the coefficient of variation (CV), expressed as a percentage, where $\%CV = (\text{standard deviation}/\text{mean}) \times 100$. Two types of variability are often discussed in the literature: Intersubject variability is the variability described between independent subjects, whereas intrasubject variability is the variability described within a subject. In the case of HVDP we are concerned with intrasubject variability. Consider a single subject who received a formulation from a drug product which is considered a HVDP in four separate sampling periods, with a washout between periods of 7 days (Figure 1).

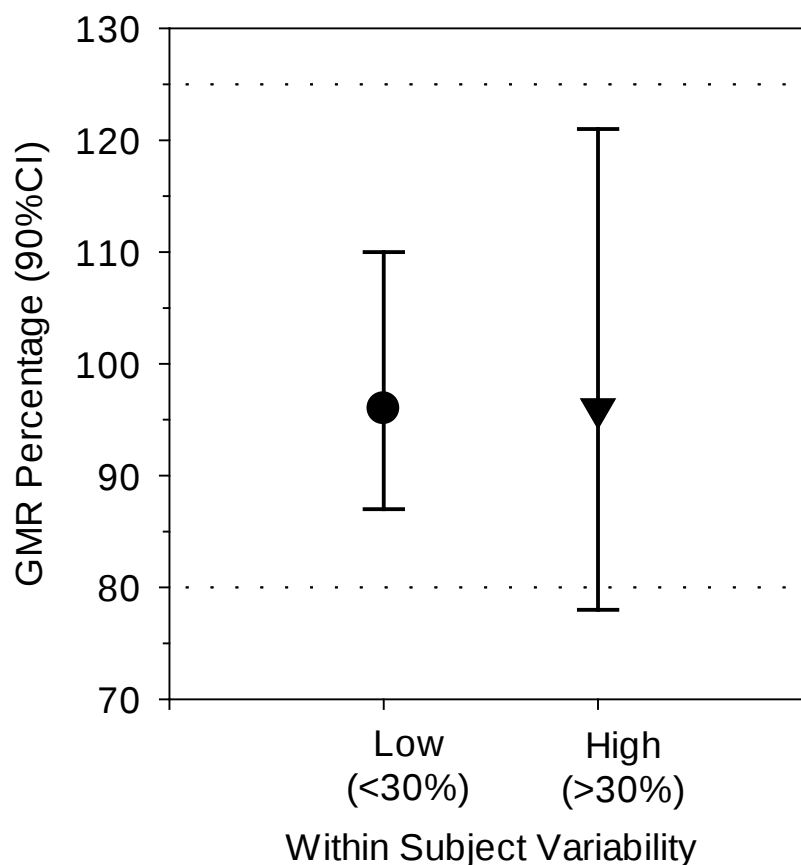
Figure 1. Plasma concentration profiles for a single subject administered 4 doses of a formulation.



In this example, note that the pharmacokinetic parameter for maximum concentration (C_{max}) ranges between 125 ng/mL and 190 ng/mL, with similar ranges of the area under the concentration curve (AUC). The intrasubject variability for this subject is around 38%, and is indicative of the overall study population for this formulation. This drug product meets the definition of a Highly Variable Drug Product (HVDP), and given this type of variability it would be difficult to establish bioequivalence with a two-period crossover study design to examine the geometric mean ratios of the test to reference products.

If one were to examine what would happen in a traditional two-period bioequivalence study design to the 90% confidence intervals in the presence of subjects with high intrasubject variability then Figure 2 provides a heuristic view. Note, in Figure 2, that the Geometric Mean ratio (GMR) is the same for low WSV and high WSV. In the presence of high WSV (intrasubject variability > 30%) the 90% confidence intervals exceeds the lower bound (<80%) and the drug product does not pass for bioequivalence.

Figure 2. Example of 90%CI with low and High Within Subject variability (WSV).



Tanguay et al (2002) noted the following failure rates from 800 published studies with increasing intrasubject %CV.

Intra-Subject CV%z	Studies Failing Bioequivalence (%)
<10%	6%
10-20%	10%
20-30%	26%
>30%	62%
Abstracted from M. Tanguay et al., AAPS Abstract, November 2002	

These failure rates clearly demonstrate that drug products with large amounts of intrasubject variability (>30%) will fail traditional bioequivalence testing. More importantly these data suggest that drug-products not previously thought to have a large WSV coefficient may in fact be classified as HVDP upon further study.

REGULATORY GUIDANCE ON HVDP

The regulatory conditions and requirements of both EMA and FDA for assessment of bioequivalence for highly variable drug products are complicated and contain various stipulations. As a result, the change of sample sizes with increasing variation is also complicated. Below are abstracted summaries from the regulatory guidance's regarding each agency's current guidance.

HEALTH CANADA GUIDANCE

The following was abstracted from the Health Canada Guidance, published on June 26, 2015 (File Number: 15-107147-494):

Previous Health Canada guidance has emphasized study design to address the challenges presented by highly variable drug products (HVDPs). Nevertheless, Health Canada's view has continued to evolve with ongoing external consultations on this issue. Health Canada recognizes that when the within-subject variation of a pharmacokinetic parameter is high, a larger number of subjects must be recruited in order to meet the usual bioequivalence standard of the 90% confidence interval within the bioequivalence interval of 80.0- 125.0%. Other regulatory agencies have also recognized the issue with HVDPs and have developed approaches to reduce the number of subjects required to meet their regulatory standards. Further to recommendations made by Health Canada's Scientific Advisory Committee on Pharmaceutical Sciences and Clinical Pharmacology, in June 2014, Health Canada has reviewed various approaches and is proposing to adopt an average bioequivalence approach to HVDP with expanding limits based on the within-subject variability of the reference product. The proposed approach would permit widening of the bioequivalence interval for AUC, with a point estimate constraint. A drug product may be considered a HVDP if the within-subject coefficient of variation (CV) of the AUC 1 for the reference product is greater than 30.0%. Critical dose drugs 2 are not eligible for the application of this approach of widening the bioequivalence intervals. Evidence from the literature, or the results of well conducted studies, should be provided to indicate that the AUC is highly variable. The proposal for widening the bioequivalence interval should be defined a priori in the study protocol. A scientific rationale should be provided to support that the high variability in exposure is not clinically significant. Submissions for HVDPs should also be supported by a justification to demonstrate that the CV estimates are reliable and not subject to the influence of outliers. For HVDPs, replicate design comparative bioavailability studies should be conducted with the reference product (R) administered at least twice to determine the within-subject variability for the reference product. The test product (T) should be administered either once in a 3-period design (RTR, TRR, RRT) or twice in a 4-period design (TRTR, RTRT). The lower and upper limits of the expanded bioequivalence interval for the 90% confidence interval should be calculated using the within-subject standard deviation of the log-transformed values of AUC of the reference product (sWR). Expansion of the 90% confidence limits may be permitted up to a maximum sWR of 0.472 (equivalent to a CV of 50.0%). For HVDPs, the following comparative bioavailability standards should be met:

1. The 90% confidence interval of the relative mean AUC of the test to reference product should be within the following limits:
 - a. 80.0%-125.0%, if $sWR \leq 0.294$ (i.e., $CV \leq 30.0\%$),
 - b. $[\exp(-0.76sWR) \times 100.0\%] - [\exp(0.76sWR) \times 100.0\%]$ if $0.294 < sWR \leq 0.472$ (i.e., $30.0\% < CV \leq 50.0\%$), or
 - c. 69.8%-143.2%, if $sWR > 0.472$ (i.e., $CV > 50.0\%$);
2. The relative mean AUC of the test to reference product should be within 80.0% and 125.0% inclusive;
3. The relative mean maximum concentration (C_{max}) of the test to reference product should be between 80.0% and 125.0% inclusive.

EUROPEAN MEDICINES AGENCY GUIDANCE

The European Medicines Agency also recognizes certain drugs as highly variable drug products (HVDP) and is willing to accept a wider difference (i.e., a wider 90% confidence interval) in C_{max} for bioequivalence evaluation. In its guidance "[Guideline on the Investigation of Bioequivalence](#)", section 4.1.10 specifically discussed the HVDP:

4.1.10 Highly variable drugs or drug products

Highly variable drug products (HVDP) are those whose intra-subject variability for a parameter is larger than 30%. If an applicant suspects that a drug product can be considered as highly variable in its rate and/or extent of absorption, a replicate cross-over design study can be carried out.

Those HVDP for which a wider difference in C_{max} is considered clinically irrelevant based on a sound clinical justification can be assessed with a widened acceptance range. If this is the case the acceptance criteria for C_{max} can be widened to a maximum of 69.84 – 143.19%. For the acceptance interval to be widened the bioequivalence study must be of a replicate design where it has been demonstrated that the within-subject variability for C_{max} of the reference compound in the study is >30%. The applicant should justify that the calculated intra-subject variability is a reliable estimate and that it is not the result of outliers. The request for widened interval must be prospectively specified in the protocol.

The extent of the widening is defined based upon the within-subject variability seen in the bioequivalence study using scaled-average-bioequivalence according to $[U, L] = \exp [\pm k \cdot s_{WR}]$, where U is the upper limit of the acceptance range, L is the lower limit of the acceptance range, k is the regulatory constant set to 0.760 and s_{WR} is the within-subject standard deviation of the log-transformed values of C_{max} of the reference product. The table below gives examples of how different levels of variability lead to different acceptance limits using this methodology.

Within-subject CV (%)*	Lower Limit	Upper Limit
30	80.00	125.00
35	77.23	129.48
40	74.62	134.02
45	72.15	138.59
≥50	69.84	143.19

$$* CV(\%) = 100\sqrt{e^{s_{WR}^2} - 1}$$

The geometric mean ratio (GMR) should lie within the conventional acceptance range 80.00-125.00%.

The possibility to widen the acceptance criteria based on high intra-subject variability does not apply to AUC where the acceptance range should remain at 80.00 – 125.00% regardless of variability.

It is acceptable to apply either a 3-period or a 4-period crossover scheme in the replicate design study.

FDA GUIDANCE

The US FDA also recognizes certain drugs as highly variable drug products (HVDP) and has stipulated in guidance that sponsors are to discuss with reviewing divisions. In the FDA Guidance Bioavailability and Bioequivalence Studies Submitted in NDAs or INDs— General Considerations (March-2014):

H. Drug Products With High Intrasubject Variability

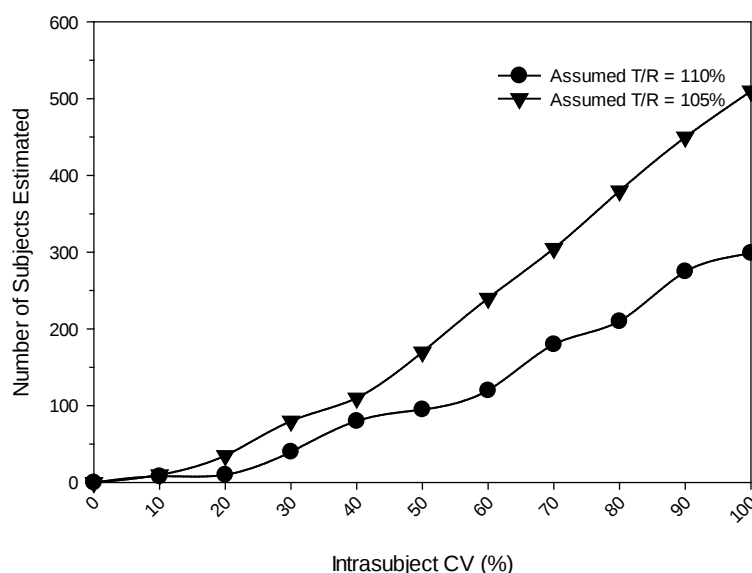
In addition to the traditional approach and the use of average BE using replicate designs, the use of a reference-scaled BE approach using a replicate design can be considered. This approach should be reserved for drugs that demonstrate a high intrasubject variability (≥30%). The reference-scaled average BE approach adjusts the BE limits of highly variable drugs by scaling to the within-subject variability of the reference product in the study and imposes a limit of 0.8 to 1.25 on the geometric mean ratio.²⁹ The appropriate review division should be consulted when planning the use of the reference-scaled BE approach.

REPLICATED CROSS-OVER STUDY DESIGNS

SAMPLE SIZE

If we were to apply the common two-period, two sequence, crossover study design with HVDPs then the sample sizes become prohibitive and more than likely be a source of ethical concerns for unnecessary exposure. Consider that with increasing intrasubject CV the number of subjects required to demonstrate bioequivalence increases exponentially (Figure 3)

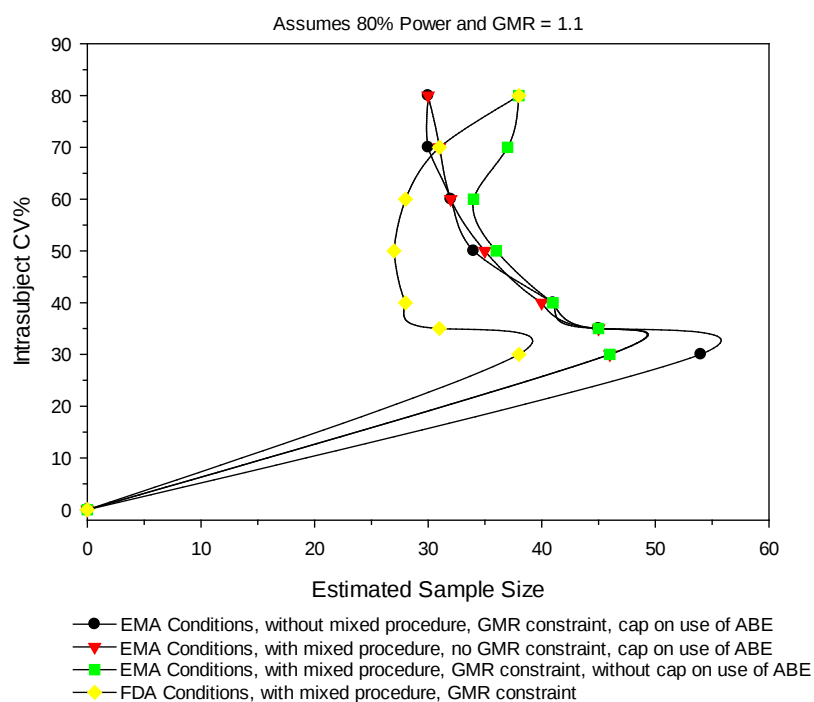
Figure 3. Estimated number of subjects for two-period, two-sequence crossover study (80% Power).



Abstracted (and recalculated) from DiLiberti (2004) presentation to FDA advisory committee for Pharmaceutical Sciences. April 14, 2004

Tothfalusi and Endrenyi (2012) discuss in a seminal manuscript the complication associated with accurately estimating sample sizes for HVDP bioequivalence studies under the varying regulatory conditions in the EMA, Health Canada, and the FDA. For example, Figure 4 presents under an assumption of 80% power and a fixed GMR estimate of 1.1 the various sample size estimates for conditions expressed by the EMA and the FDA.

Figure 4. Sample size estimates under regulatory guidance conditions [From Tothfalusi and Endrenyi (2012)]



REPLICATED STUDY DESIGNS AND HVDP

The use of replicated (partially or fully) replicated study designs for HVDP is desirable to reduce the number of subjects required for the study and increase study power. Several advantages to replicated designs are noted as follows:

- Allows for a comprehensive evaluation of the intrasubject variances for both the test and reference products,
- Provides valuable information on whether a test product demonstrates higher or lower intrasubject variability in the bioavailability measures compared to the reference product.
- Provides valuable information on the intrasubject variability of the reference product (intra- or inter-formulation variability, as well as within a subject variability).
- Information on the intrinsic factors associated with the formulation manufacturing processes.
- Replicated study designs will always be associated with smaller numbers of subjects in a study.
- Replicated designs increase study power when intrasubject variability in systemic exposure of the test (and reference) drug products is high (>30%).

Among various approaches to address the bioequivalence issue for highly variable drugs, reference-scaled average bioequivalence (ABE) has been suggested in the literature. This approach requires less subjects in the study, but with replicated treatment design. The replicated crossover designs were also discussed in selected FDA guidance (Statistical Approaches to Establish Bioequivalence), but this discussion was for dealing with the carryover effects. In this discussion, the replicated crossover designs are for dealing with highly variable drugs and the intra-subject variability. We summarize the use of the three-period (partially replicated) and 4-period (fully-replicated) study designs for HVDP BE studies. It should be noted the increased level of information obtained from the more sophisticated replicated study designs. For instance the more sophisticated the design the better the estimate of overall variance:

- Hierarchy of BE study designs:
 1. Full Replicate (e.g. TRTR|RTRT or TRT|RTTR)
 2. Partial Replicate (e.g. TRR|RTR|RRT)
 3. 2x2 Crossover (TR|RT)
 4. Parallel (T|R)
- More Robust estimates of Variance from BE Study Designs:
 1. Full Replicate: Total Variance (between + within subjects) + within subjects (R and T)
 2. Partial replicate: Total Variance (between + within subjects) + within subjects (R)
 3. 2x2 Crossover: Total Variance (between + within subjects) + between, within subjects
 4. Parallel: Total Variance (between + within subjects)

Replicated Bioequivalence study designs may be depicted in the following table.

Design	Sequence	Study Period			
		I	II	III	IV
Partially-Replicated, 3-period, 3-sequence, Reference product is replicated	TRR	T	R	R	N/A
	RTR	R	T	R	
	RRT	R	R	T	
Partially-Replicated, 3-period, 3-sequence, Test product is replicated	RTT	R	T	T	N/A
	TRT	T	R	T	
	TTR	T	T	R	
Fully-Replicated, 4-period, 6-sequence	RTTR	R	T	T	R
	TRTR	T	R	T	R
	TTRR	T	T	R	R
	RTRT	R	T	R	T
	RRTT	R	R	T	T
	TRRT	T	R	R	T
Fully-Replicated, 4-period, 2-sequence	RTRT	R	T	R	T
	TRTR	T	R	T	R

T = Test Product, R = Reference Product

The choice of a fully replicated 6-sequence or 2-sequence design is dependent upon sample size and regulatory requirements. Often times a 2-sequence, 4-period, fully replicated design is sufficient to provide an assessment of within subject's variability (intrasubject variability).

CONCLUSIONS

The assessment of bioequivalence between reference and test products has scientifically sound and reproducible criteria for acceptance of a drug product as a generic equivalent with the 80-125% boundaries on the 90% confidence intervals of the geometric mean ratio for the pharmacokinetic parameters C_{max} and AUC. These boundaries are universally accepted with drug products that present WSV (intrasubject variation) of <30%. These boundaries are difficult to reproduce for HVDP where the WSV is >30%, and the regulatory agencies have been wrestling with these issues for over 10 years. Because of the complexity in study design, sample size, and model selection with HVDP, it should be concluded that sponsors must, before conducting bioequivalence studies with HVDP candidates, consult with appropriate regulatory authorities in the EMA, FDA, and Health Canada.

The literature and data to date suggest that fully replicated study designs (4-period, 2- or 6-sequence randomization) with mixed models accounting for within subject variation, and a subject by formulation, should be considered as the most robust assessment for examining bioequivalence with HVDP. Wherever possible the use of ABE should be considered in the assessment of bioequivalence. We cannot recommend using a partially replicated study design unless the known half-life of the drug product under examination is so long that washout between study periods leads to loss of subjects in the study.

Also given the complexity of HVDP a small pilot study to adequately characterize the WSV is desirable for planning the definitive bioequivalence study. This pilot study can also be used to fully explore model selection criteria and appropriate GMR ratios that are meaningful in planning the appropriate sample size.

REFERENCES AND RECOMMENDED READING

- Haidar SH, Davit B, Chen ML, Conner D, Lee L Li QH, Lionberger R, Makhoul F, Patel D, Schuirmann DJ, Yu LX (2008). Bioequivalence approaches for highly variable drugs and drug products. *Pharmaceutical research* January-2008; 25(1):237-241. Epub 2007 September 22.
- Midha KK, Rawson MJ, Hubbard JW (2005). The bioequivalence of highly variable drugs and drug products. *International Journal of Clinical Pharmacology and Therapeutics*. October-2005; 43(10):485-498.
- Tothfalusi L and Endrenyi L (2012) Sample Sizes for Designing Bioequivalence Studies for Highly Variable Drugs. *J Pharm Biopharm Sci* 15(1) 73-84.
- Guidance for Industry, Statistical Approaches to Establishing Bioequivalence. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). January 2001.
- Jones B and Kenward MG (2003). *Design and Analysis of Cross-Over Trials*, Second Edition. Chapman & Hall/CRC, New York.
- Machado S, Miller R and Hu C (1999). A Regulatory Perspective on Pharmacokinetic and Pharmacodynamic Modelling. *Statistical Methods in Medical Research*, 8: 217-245.
- Mohsen MA (2010) A Note on the Calculation of Intrasubject Coefficient of Variation in Bioequivalence Trials. *J Bioanal Biomed* 2: 075-078. doi:10.4172/1948-593X.1000026
- Patterson SD and Jones B (2002a). Bioequivalence and the Pharmaceutical Industry. *Pharmaceutical Statistics*, 1:83-95.
- Patterson SD and Jones B (2002b). Replicated Designs and Average, Individual, and Population Bioequivalence. GlaxoSmithKline BDS Technical Report 2002-05, 11 December 2002.
- Schuirmann DJ (1981). On hypothesis testing to determine if the mean of a normal distribution is contained in a known interval. *Biometrics* 37:617.
- Schuirmann DJ (1987). A comparison of the two one sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J. Pharmacokin. Biopharm.* 15: 657-680.
- Richardson BA, Flack VF (1996). The analysis of incomplete data in the three-period two treatment cross-over design for clinical trials. *Statistics in Medicine*. 15(2), 127-143.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author via email at:

James R. Johnson, PhD (Jim)
Cary, NC USA
Email: jrjphd@gmail.com

Brand and product names are trademarks of their respective companies.