

c

DECLARATION OF WALTER W. HAUCK, PH.D.

In Support of the Citizen Petition of Abbott Laboratories Docket No. 2003P-0387/CP1

Walter W. Hauck, Ph.D., under penalty of perjury, declares as follows:

1. Abbott Laboratories ("Abbott") has requested that I comment on the Food and Drug Administration's ("FDA's") standard statistical methodology for analyzing bioequivalence data when applied to levothyroxine sodium drug products.
2. For this declaration, I have reviewed, among other materials, Abbott's Citizen Petition (including attachments), several supplements and comments to the Citizen Petition docket, the full report (including appendices) of Abbott's Study M02-417, the preface to FDA's publication, *Approved Drug Products with Therapeutic Equivalence Evaluations* (also known as the "Orange Book"), and FDA's guidance documents on general bioavailability/bioequivalence testing and statistical approaches to establishing bioequivalence.

Qualifications

3. I am Professor of Medicine and Head of the Biostatistics Section, Division of Clinical Pharmacology, Department of Medicine at Thomas Jefferson University. I also am Director of Biostatistics at the Kimmel Cancer Center at Thomas Jefferson University. My *curriculum vitae* is attached at Tab 1.
4. I hold a Ph.D. and a Master's degree in statistics from Harvard University. My Bachelor's degree (mathematics and economics) is from Carnegie-Mellon University.

5. My expertise is in the biostatistical aspects of clinical pharmacology, particularly bioequivalence studies. Currently, my statistical research focuses on methods for demonstrating equivalence, including the equivalence of generic pharmaceutical products. I was the primary investigator on a contract from FDA that supported a portion of this work.

6. I have authored or co-authored nearly 200 articles and book chapters, many dealing with bioequivalence issues. I also have delivered many invited presentations, including many on the concepts of average, individual, and population bioequivalence.

7. I have served as a consultant to FDA's Office of Pharmaceutical Science and the United States Pharmacopeia. I also serve, or have served, as a referee or on the editorial boards of numerous journals, including *Statistics in Medicine*, *Journal of Pharmacokinetics and Pharmacodynamics*, and *International Journal of Clinical Pharmacology and Therapeutics*.

Bioequivalence Testing

8. Bioavailability is defined as the rate and extent to which the active ingredient in a drug product is absorbed and becomes available at the site of action.

9. Bioequivalence is essentially a determination that two drug products (a "test" product and a "reference" product), when administered in the same amounts under the same experimental conditions, do not differ in their relative bioavailability by a clinically significant amount.

10. According to FDA, pharmaceutically equivalent products (*i.e.*, those containing identical amounts of the identical active ingredient in the identical strength and dosage form) that

are found to be bioequivalent may be substituted for one another with the full expectation that they will have the same clinical effect and safety profile, without any adjustment in dose or additional therapeutic monitoring.

11. Bioequivalence studies typically are conducted as single-dose, two-treatment, crossover studies. A single dose of either the test or reference product is administered to each subject, and blood samples are taken repeatedly over an appropriate period of time. After an adequate interval (called a “washout period”), during which the first dose is eliminated from each subject’s body, the other product (test or reference) is administered, and the sampling is repeated.

12. The comparison of the test and reference products within each individual subject ensures that differences in the way different individuals absorb the drug (called “inter-subject variability”) do not impact the study. This also allows for estimation of differences in the way the same individual absorbs the drug after the two doses (called “intra-subject variability”).

13. After both products have been administered to each subject, pharmacokinetic measures characterizing the rate and extent of absorption of the drug are developed. Typically, for solid oral drug products, FDA recommends that bioequivalence studies measure the peak drug concentration (“ C_{max} ”) and the area under the concentration-time curve (“AUC”).

14. C_{max} represents the maximum concentration of the drug in the body. AUC is a measure that is proportional to the total amount of drug reaching the systemic circulation and thus characterizes the extent of absorption. “ T_{max} ” is another pharmacokinetic measure that is

sometimes used in bioequivalence testing, and is defined as the time that it takes for the drug to reach C_{\max} .

15. AUC measurements typically are developed for several different times throughout the measurement period. For example, $AUC_{48 \text{ hr}}$ represents the area under the concentration-time curve for the first 48 hours of a study.

FDA's Statistical Methodology

16. FDA recommends that individual C_{\max} and AUC measurements be logarithmically transformed. The logarithmic transformation has two motivations. First, because AUC is proportional to bioavailability, the logarithm converts a multiplicative relationship into a linear one. Statistical models are primarily for linear relationships. Results can then be transformed back to the original scale, and expressed in terms of a test-to-reference ratio. Second, experience has been that the assumption of a normal distribution for the within-subject errors is more reasonable following logarithmic transformation.

17. Average C_{\max} and AUC measurements are developed, and the relationship of the test product's measurements to the reference product's measurements is expressed as a ratio of those averages. An analysis of variance ("ANOVA") is then performed to develop a 90% confidence interval for the test-to-reference ratio. This interval or range of values provides 90% confidence that the "true" ratio between the test and reference products lies within that range.

18. For example, suppose the 90% confidence interval for the test-to-reference ratio of mean AUC measurements is 0.95 to 1.15. While we do not know the true ratio, we are 90% confident that the true ratio of the mean AUC of the test product to that of the reference product

lies between 0.95 and 1.15. By confidence, we mean that if we repeat the study many times, the confidence intervals will include the true ratio 90% of the time.

19. Finally, for bioequivalence purposes, FDA applies an acceptance range to the 90% confidence intervals surrounding the ratios of the mean C_{max} and AUC measurements. In developing this acceptance range, FDA appears to have made a general clinical judgment that the mean C_{max} and AUC measurements of test and reference products should not differ by more than 20%.

20. Therefore, FDA's acceptance range targets a test-to-reference ratio of less than 0.80 (preventing the mean C_{max} or AUC of the test product from being more than 20% less than the mean C_{max} or AUC of the reference product). The acceptance range likewise targets a reference-to-test ratio of less than 0.80 (preventing the mean C_{max} or AUC of the reference product from being more than 20% less than the mean C_{max} or AUC of the test product). By convention, all bioequivalence data is expressed in terms of the test-to-reference ratio, so the latter acceptance limit is expressed as 1.25, the reciprocal of 0.80. Thus, FDA's standard bioequivalence acceptance range is 0.80 to 1.25, or 80 to 125%.

21. In other words, in order to be declared bioequivalent, the 90% confidence intervals around the test-to-reference ratios for the mean C_{max} and mean AUC measurements should fall fully within 0.80 to 1.25. If either confidence interval extends outside of the 0.80 to 1.25 boundaries, in either direction, there cannot be 90% confidence that the mean C_{max} or AUC of the test product is within 20% of the reference product. FDA generally considers such products as having failed to demonstrate bioequivalence.

22. I understand that FDA chose this acceptance range based on a clinical judgment that the bioavailability of two interchangeable products should not differ by more than 20%. By setting an acceptance range of 80 to 125%, FDA is saying that any true test-to-reference ratio that falls within that range is acceptable, as long as the sponsor is able to so demonstrate. FDA's standard acceptance range is not designed to prevent differences of less than 20%. Thus, it is not designed for certain classes of drugs that must be maintained within a narrower range. For example, if the bioavailability of two products must not differ by more than 10%, due to efficacy and/or safety considerations, the acceptance range would need to be adjusted to 0.90 to 1.11, in order to ensure that only products differing by less than 10% would pass as bioequivalent.

The Impact of Sample Size and Variability

23. The width of each 90% confidence interval impacts the determination of bioequivalence. In particular, the narrower the confidence interval, the farther the ratio at the center of that interval may drift from 1.00 (perfect unity between the test and reference products), and still result in a declaration of bioequivalence. Conversely, the wider each confidence interval, the closer that ratio must be to 1.00 for the entire interval to fit within FDA's 0.80 to 1.25 acceptance range.

24. Generally, the width of each confidence interval is a function of the number of subjects tested and the amount of intra-subject variability in the data.

25. The more closely-matched each individual subject's measurements between the two treatment periods, the less intra-subject variability in the data. This narrows the resulting

confidence intervals, because there can be greater confidence that the “true” ratios between the test and reference products fall within a narrower range.

26. Table 1 shows the largest mean test-to-reference ratios that can pass the 80 to 125% acceptance range, depending on sample size and intra-subject variability. The standard deviations shown are for logarithmically-transformed data. A standard deviation in the natural log scale is approximately the coefficient of variation (“CV”) in the original scale. For example, a standard deviation of 0.20 corresponds approximately to a CV of 20%.

27. The range of standard deviations shown in Table 1 is an appropriate range for levothyroxine products, given the variabilities demonstrated in Abbott’s Studies M02-417 and M01-323 (the latter reported in the simulation study conducted by Thomas Ludden, Ph.D.), following application of Correction Method 1 to $AUC_{48 \text{ hr}}$ measurements. It is my understanding that Correction Method 1 in Study M02-417 is similar to the one that has now been adopted by FDA.

Table 1

Largest Test-To-Reference Ratio (%) That Can Pass The Current FDA Bioequivalence Acceptance Range – Dependency On Sample Size And Study’s Intra-Subject Standard Deviation

Study’s Sample Size	Study’s Intra- Subject Standard Deviation							
	0.1	0.125	0.15	0.175	0.2	0.225	0.25	0.275
8	113.4%	110.7%	108.0%	105.5%	102.9%	100.5%	None	None
12	116.1%	114.0%	111.9%	109.8%	107.8%	105.8%	103.9%	102.0%
16	117.5%	115.6%	113.9%	112.1%	110.4%	108.7%	107.0%	105.3%
20	118.3%	116.7%	115.1%	113.6%	112.0%	110.5%	109.0%	107.5%
26	119.2%	117.8%	116.4%	115.0%	113.7%	112.3%	111.0%	109.7%
33	119.9%	118.6%	117.4%	116.2%	115.0%	113.8%	112.6%	111.4%
40	120.4%	119.2%	118.1%	117.0%	115.9%	114.8%	113.8%	112.7%
54	121.0%	120.1%	119.1%	118.1%	117.2%	116.3%	115.3%	114.4%
70	121.5%	120.7%	119.8%	119.0%	118.1%	117.3%	116.5%	115.7%

28. We see from Table 1 that the larger the sample size and/or the smaller the intra-subject variability, the larger the test-to-reference ratio that can pass bioequivalence.

29. The highlighted test-to-reference ratios shown in Table 1, above, correspond to those that would be expected to pass, given typical sample sizes and the variability exhibited by levothyroxine products. For example, the intra-subject variance for AUC₄₈ from Abbott’s Study M02-417 was 0.03928484. This corresponds to a 95% confidence interval for the actual intra-subject standard deviation of 0.1671 – 0.2436, or a CV of approximately 17-24%. (Dr. Ludden reports an intra-subject CV of 21% for Study M01-323, consistent with the results of Study M02-417.)

30. Given Study M02-417’s sample size of 33 subjects, a test-to-reference ratio of 115.0% (indicating a mean test AUC measurement that is 15% higher than the mean reference

AUC measurement) could have passed as bioequivalent. Similarly, a mean reference AUC measurement that is 15% higher than the mean test AUC measurement could have passed. Based on my understanding of the clinical discussion in Abbott's Petition, FDA's standard statistical methodology is thus not calibrated to the clinical needs of levothyroxine patients.

Conclusion

31. Any of several steps might be taken to improve the sensitivity of FDA's standard statistical methodology, for application to levothyroxine sodium drug products. As noted above, FDA's acceptance range may be narrowed, in order to ensure that only products differing by a certain percentage, or less, will pass as bioequivalent. The precise range to be used should be validated with respect to the clinical significance of any differences that may occur when one product is represented to be interchangeable with another product.

32. Another option to consider is the use of greater confidence intervals, such as 95%. The impact of a 95% confidence interval instead of a 90% interval is to raise sample size requirements by about 40%. That is, the differences shown in Table 1 can still pass as bioequivalent with 95% confidence intervals, if the sponsor enters about 40% more subjects into the bioequivalence study.

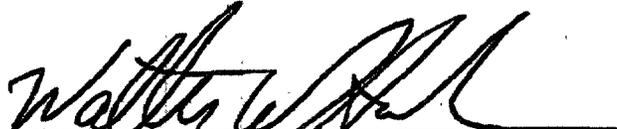
33. The key to designing an accurate bioequivalence methodology for any particular class of drugs is to begin with a clinical determination of the maximum permissible difference between products that will be regarded as interchangeable. After this determination is made, the proper bioequivalence study design, confidence interval, and acceptance range can be developed.

Declaration of Walter W. Hauck, Ph.D.

Docket No. 2003P-0387/CP1

I declare under penalty of perjury that the foregoing is true and correct.

Executed on this 13th day of April, 2004.


Walter W. Hauck, Ph.D.