HOGAN & HARTSON

L.L.P.

DAVID M. FOX
PARTNER
(202) 637-5678
DMFOX@HHLAW.COM

COLUMBIA SQUARE
555 THIRTEENTH STREET, NW
WASHINGTON, DC 20004-1109
TEL (202) 637-5600
FAX (202) 637-5910
WWW.HHLAW.COM

April 15, 2004

2004 APR 15 P3:46

BY HAND DELIVERY

Division of Dockets Management
Food and Drug Administration
5630 Fishers Lane, Room 1061
Rockville, Maryland 20852

Re: Docket No. 03P-0387
Supplement to Citizen Petition

Dear Sir or Madam:

On behalf of Abbott Laboratories ("Abbott"), we submit the following supplement under 21 CFR 10.30(g) to the above-referenced Citizen Petition, filed on August 25, 2003 (the "Petition"). Through this submission, we are:

- Adding to the record a just-published, peer-reviewed article, *Are Bioequivalence Studies of Levothyroxine Sodium Formulations in Euthyroid Volunteers Reliable?*, and an accompanying editorial, from the journal *Thyroid* (Tabs A and B);
- Adding to the record a declaration by Walter W. Hauck, Ph.D., Head of the Biostatistics Section, Department of Medicine, at Thomas Jefferson University, and Dr. Hauck's *curriculum vitae* (Tabs C and C.1);
- Responding to the February 24 and March 16, 2004, comments of Hyman, Phelps & McNamara, P.C. ("Hyman, Phelps"); and
- Providing an analysis of recent correspondence between the Food and Drug Administration ("FDA") and the American Thyroid Association ("ATA"), The Endocrine Society, and the American Association of Clinical Endocrinologists (the "Medical Societies") with regard to a workshop on levothyroxine bioequivalence ("BE") testing (Tab D).

2003P-0387 SUP 6

I. PUBLICATION OF A PEER-REVIEWED ARTICLE AND EDITORIAL ON STUDY M02-417

We are submitting to the record a copy of a peer-reviewed article, *Are Bioequivalence Studies of Levothyroxine Sodium Formulations In Euthyroid Volunteers Reliable?*, and an accompanying editorial, both of which were published in the March 2004 issue of *Thyroid*, the official journal of the ATA. *See* Tabs A and B. The article, authored by Vicky Blakesley, M.D., Ph.D., *et al.*, describes Abbott's Study M02-417, a primary piece of evidence in support of the Petition.

In the editorial, Jim Stockigt, M.D., introduces the article and the "simple premise" of Abbott's clinical study: "[A] test that shows no significant difference between two different doses would be wrongly interpreted as demonstrating bioequivalence." Tab B at 2.[1] Dr. Stockigt discusses how the use of FDA's methodology to establish equivalence is problematic for an endogenous, narrow therapeutic index drug like levothyroxine, particularly where there is evidence to show that subclinical thyrotoxicosis and mild thyroid failure can have important biological consequences. *See id.*

The article itself provides clinicians with detailed information on the design of Study M02-417, along with the study's key conclusions. It describes how, without correction for baseline levels of endogenous thyroxine, FDA's previously recommended BE methodology could not detect differences between products of up to 33%. Even with any of three different methods of baseline correction – including the method eventually adopted by FDA – differences of 12.5% could not be detected. *See* Tab A at 6-7. As a result, the article explains that baseline correction is a necessary, but by no means sufficient, step in refining FDA's methodology as it is applied to levothyroxine. The article concludes that still "[m]ore precise methods for defining bioequivalence are required in order to ensure that [levothyroxine] products accepted as bioequivalent will perform equivalently in patients without the need for further monitoring and retitration of their dose." *Id.* at 1.

In short, as Dr. Stockigt explains, Blakesley, *et al.*, showed that FDA's recommended methodology could not detect a clear and clinically relevant dosing difference. Of FDA's methodology: "If a technique is to be used to identify small differences, it is obviously relevant to know what differences can be detected, *but this does not seem to have been established.*" Tab B at 1 (emphasis added).

---

1 Abbott's understanding is that Dr. Stockigt was selected by the editors of *Thyroid* because he has no financial relationship with any companies whose products might be affected by this issue.

renders use of FDA's standard 80-125% acceptance range clinically inappropriate. Comment at 1.[2]

As shown below, this Comment does more to support Abbott's thesis than it does to undermine it. Most significantly, Hyman, Phelps plainly agrees with Abbott's core statistical argument – namely, that the application of FDA's standard statistical methodology to levothyroxine products can result in a declaration of bioequivalence for products that differ by a clinically significant amount. *See infra* at III.B.

A. TSH as an Additional Marker

The first section of the Comment is devoted to rebutting a point not argued by Abbott. Hyman, Phelps states that Abbott's Petition argues that it is necessary to measure TSH for BE purposes. *See* Comment at 2-4. Nowhere does Abbott's Petition or its supplements state that it is necessary to measure TSH rather than total thyroxine in BE testing, and Hyman, Phelps provides no citations for its claim. Rather, Abbott included in a list of issues to be presented to a joint advisory committee the possibility of "*additional markers* to assess the bioequivalence of levothyroxine, such as TSH – particularly in light of clinicians' reliance on TSH in titrating patients to specific doses of levothyroxine." Petition at 41 (emphasis added).

That being said, Abbott respects the views of the numerous clinical experts who believe that TSH measurements should play a role in FDA's BE determinations. Because these clinicians use TSH to adjust their patients' doses of levothyroxine, many reason that products approved by FDA as interchangeable must produce comparable TSH levels. Moreover, there is precedent for the use of such secondary BE measures. For example, FDA currently evaluates "T_{max}" – the time to maximum serum concentration – in BE studies, but does not apply any fixed acceptance range. *See* Guidance for Industry: *Bioavailability and Bioequivalence Studies for Orally Administered Drug Products – General Considerations* 22 (March 2003) ("General BA/BE Guidance"). Again, while Abbott does not recommend the use of TSH as a primary marker for equivalence, we do believe that the strongly held opinions of the leading endocrinologists must be taken into account. *See, e.g.*, Tab B at 2 ("It is well established and widely accepted that the bioavailability of

---

2 On March 16, 2004, Hyman, Phelps supplemented its Comment to correct several misstatements in its earlier submission and to provide additional information documenting the low variability of levothyroxine measurements. *See also* Comment at 3, 4, 5.

thyroxine preparations is best assessed by detailed examination of the decline in serum TSH.") (footnotes omitted).

B. Inadequacy of FDA's Statistical Methodology

Hyman, Phelps asserts that FDA's currently recommended BE methodology for levothyroxine is "scientifically sound and appropriate." Comment at 4. However, the Comment itself acknowledges that levothyroxine products that differ in dose by a significant amount will be declared equivalent under FDA's methodology: "Therefore, *products that are 12.5% different will pass the current confidence interval guidelines*, using an adequate sample size." Comment at 5 (emphasis added). As discussed in Abbott's Petition and in four declarations from leading endocrinologists, differences in dose of this amount present significant risk to patients. *See* Petition at 4-5, 22-28; Supplement to Petition (Feb. 9, 2004).

Nevertheless, Hyman, Phelps argues that it is "extremely unlikely" that FDA will be asked to find such products bioequivalent. Comment at 5. This is so, according to the Comment, because "FDA also evaluates other data, in addition to bioequivalence studies, to assure the products will perform comparably to a reference listed drug. For example, FDA considers the dissolution profile and the formula as part of its global evaluation." *Id.* "All of this," Hyman, Phelps concludes, "helps to assure that products that meet the current FDA guidelines will behave the same as the reference drug." *Id.*

In other words, Hyman, Phelps argues that, with potency and dissolution data, the agency can salvage an *in vivo* BE methodology that cannot distinguish among products known to deliver different amounts of levothyroxine to the body. As discussed below, however, *in vitro* measures cannot compensate for the deficiencies inherent in FDA's recommended BE methodology.

1. Limitations of In Vitro *Analysis*

The *in vivo* demonstration of equivalence is the recognized standard for products intended to be absorbed into the bloodstream, and is regarded as more sensitive than *in vitro* testing. *See* 21 USC 355(j)(8)(B); 21 CFR 320.24(b).[3] This

---

3 Generally, *in vitro* determinations of equivalence are sufficient only when a product is in the same dosage form, but in a different strength, and is proportionately similar in active and inactive ingredients as a product for which the sponsor has demonstrated *in vivo* equivalence; when there has been an *in vitro/in vivo* correlation; or when there has been a reformulation that could not affect bioavailability. *See* 21 CFR 320.22(d). *In vitro/in vivo* correlations can be used for extended release drug products, but are formulation- and manufacturer-specific. They are not recommended for the approval of different sponsors' products. *See* Guidance for Industry: *Extended Release Oral Dosage*

need for *in vivo* BE testing for solid oral drug products is based on FDA's sound judgment that *in vitro* analysis cannot assure equivalence inside the body. Even when two sponsors' products are pharmaceutically equivalent (*i.e.*, they contain identical amounts of the same active ingredient in the same dosage form), they may release different amounts of drug at different rates, resulting in variable absorption. As noted above, for a narrow therapeutic index drug like levothyroxine, even small differences in dose may have clinically significant effects on patients.[4] *See generally* General BA/BE Guidance at 20 (recommending *additional* testing of narrow therapeutic index drugs to assure interchangeability).

The limitations of *in vitro* testing have been outlined fully by FDA, most clearly in a presentation by Ajaz S. Hussain, Ph.D., to the Advisory Committee for Pharmaceutical Science ("ACPS"). As explained by Dr. Hussain (then an acting director in the Office of Pharmaceutical Science), particle size, excipients, manufacturing process, equipment, and batch size can be different for pharmaceutically equivalent products, and each can affect *in vivo* bioavailability. Transcript (Nov. 16, 2000) at 16, 18 at www.fda.gov/ohrms/dockets/ac/cder00.htm. The media composition and volume, and hydrodynamics, of dissolution tests may also not be comparable to the *in vivo* environment. Finally, gastric residence time, surface tension, and the permeability of the small intestine are likewise among the "uncertainties and complexities that," according to Dr. Hussain, "are not captured in *in vitro* dissolution." *Id.* at 19-22. For these reasons, such testing is used for

---

Forms: Development, Evaluation, and Application of In Vitro/In Vivo *Correlations* 16 (Sept. 1997). And, sponsors may request waivers from *in vivo* BE testing for rapidly dissolving immediate release drug products containing highly soluble and highly permeable drug substances. *See* Guidance for Industry: *Waiver of* In Vivo *Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System* ("BCS") (Aug. 2000). Levothyroxine is not such a drug and, in any case, BCS-based biowaivers are not available for narrow therapeutic index drugs. *See id.* at 9.

4 Hyman, Phelps asserts that there is no "formal designation" for narrow therapeutic index drugs. Comment at 6. Formality aside, FDA clearly recognizes such drugs in numerous places. *See, e.g.*, 21 CFR 320.33(c); General BA/BE Guidance at 20; and Guidance for Industry: *Immediate Release Solid Oral Dosage Forms: Scale-Up and Postapproval Changes: Chemistry, Manufacturing, and Controls,* In Vitro *Dissolution Testing, and* In Vivo *Bioequivalence Documentation* Appendix A (Nov. 1995). FDA also has referred specifically to levothyroxine as a narrow therapeutic index drug, most notably in the approved labeling for this class of products. *See* Petition, Tab 7, at 262; *see also* 62 FR 43535, 43538 (Aug. 14, 1997); *Guidance for Industry: Levothyroxine Sodium Tablets* – In Vivo *Pharmacokinetic and Bioavailability Studies and* In Vitro *Dissolution Testing* 2 (Dec. 2000). Even the first attachment to the Hyman, Phelps Comment states, "FDA's . . . classification of this drug as a narrow therapeutic range drug was also taken into consideration." Comment, Tab 1, at 3.

quality control, but may not act as a general substitute for *in vivo* BE testing. *See* 21 CFR 320.22(d).

Simply put, no amount of potency and dissolution testing can assure bioequivalence when the applicable *in vivo* methodology is itself incapable of distinguishing among products known to deliver different amounts of the drug.

2. *Two Illustrations of the Need for* In Vivo *Analysis of Levothyroxine Products*

In requesting that FDA rely on *in vitro* potency and dissolution testing, Hyman, Phelps ignores significant evidence that such testing is *not* an accurate predictor of *in vivo* levothyroxine bioavailability. The bioavailability studies submitted to FDA in support of the approved levothyroxine drug products provide several examples.

The first example is found in FDA's approval of Mylan Laboratories' generic version of Unithroid®. In commenting on Mylan's BE studies, FDA reports the results of its application of two methods of baseline correction. *See* Petition, Tab 11, at 477. In a BE study comparing two 600 mcg doses of levothyroxine, the test product was found to be 8% less bioavailable than the reference product, *in vivo*. *See id.* at 479. After application of Correction Method 1, the test-to-reference ratio for the mean area under the serum concentration-time curve ("AUC") was only 0.92, with a 90% confidence interval ("CI") of 0.85-0.99. The corrected ratio for the maximum serum concentration ("C_{max}") was lower, or 0.91 (90% CI 0.86-0.97). The test product demonstrated this lower *in vivo* bioavailability despite the fact that its potency tested nearly 5% *higher* than the potency of the reference product. *See id.* at 373. The *in vitro* demonstration clearly did not translate *in vivo*.

Likewise, at the March 13, 2003, ACPS meeting, Steven B. Johnson, Pharm.D., speaking for FDA, presented a slide showing dosage form proportionality data for four brand name levothyroxine products. *See* Petition at 30. The slide was intended to demonstrate the sensitivity of FDA's baseline correction method. In fact, the slide shows how *in vitro* measures do not assure the *in vivo* equivalence of levothyroxine products. According to Dr. Johnson, when corrected for baseline, two 600 mcg doses of a single sponsor's levothyroxine product differed *in vivo* by 14%. Two 600 mcg doses of another sponsor's product differed *in vivo* by 8.5%. *See id.* In both cases, the different configurations of each sponsor's product likely contained the same formulation, and met the same dissolution specification. Again, the *in vitro* analysis of levothyroxine products cannot assure that the results of Study M02-417 will not be borne out in future BE studies.

Hyman, Phelps concludes its Comment by arguing that, "[i]t would not be surprising to see differences of average potency being close to 10% between production and prolonged storage. In addition, individual tablets vary in their potency and dissolution." Comment at 6. Of course, any variation in potency due to storage would be *cumulative* to any variation attributable to FDA's faulty BE methodology. Even if Hyman, Phelps' assertions about prolonged storage are accurate, this merely presents an additional reason to ensure that "bioequivalent" levothyroxine products are closely matched at their initial release.

Hyman, Phelps thus offers no evidence to counter the fact that the agency's standard BE methodology cannot distinguish levothyroxine products that differ by a significant amount. Indeed, the Comment concedes that products that are 12.5% different will pass the current criteria. No amount of *in vitro* testing can compensate for the deficiencies in FDA's methodology.

IV. CORRESPONDENCE REGARDING LEVOTHYROXINE BIOEQUIVALENCE WORKSHOP

A. Background

On October 28, 2003, the Medical Societies submitted to Docket No. 03P-0387 the text of a letter from FDA, describing the agency's decision to hold a workshop on the issue of BE standards for levothyroxine products. The letter described FDA's commitment to "hold a workshop of sufficient depth and duration to address all of the relevant issues," including baseline correction, optimal test subjects, acceptable confidence limits, and TSH. Tab D at 1.

In light of this positive development, Abbott supplemented its Petition to state that this scientific workshop, appropriately structured, would satisfy Abbott's request for an advisory committee or similar scientific meeting. *See* Supplement to Petition (Dec. 22, 2003). Shortly thereafter, and at FDA's request, the Medical Societies submitted to the agency a draft agenda for the workshop. *See* Supplement to Petition (Jan. 9, 2004).

We now understand that the agency has decided to postpone all planning for the workshop. According to a letter from the Medical Societies to FDA dated March 19, 2004, FDA has decided to postpone the meeting because the agency wants first to resolve the issues raised by the pending levothyroxine citizen petitions. *See* Tab D at 1. By postponing the meeting, FDA is depriving itself of the relevant information needed to craft a sufficient response. The agency's initial basis for holding the workshop – *i.e.* to hear from the leading experts – makes its apparent decision to indefinitely postpone the meeting all the more troubling.

B. The Agency's Pattern of Conduct Calls Into Question the Sufficiency of the Process and the Administrative Record

This recent conduct continues a troubling pattern of FDA refusing to convene a meaningful meeting on levothyroxine BE testing, or to engage in any sort of useful dialogue on the issue. As such, should FDA refuse to grant Abbott the relief sought in the Petition, the agency's action will not be granted deference by the courts. The following actions illustrate the agency's troubling pattern of refusing to meet with experts or to consider relevant factors:

- **Without adequate explanation, FDA denied Abbott's repeated requests to meet before the agency made its initial decision to adopt baseline correction**: In 2002, FDA refused to meet with Abbott on the design of Study M02-417 and then, subsequent to completion of the study, denied Abbott's renewed request for a meeting. In a letter to Abbott rejecting the second meeting request, FDA provided by way of explanation its adoption of "a three pre-dose baseline subtraction method to evaluate total thyroxine" when considering products for "AB" therapeutic equivalence ratings. Petition, Tab 4, at 1. FDA offered no discussion of the data from Study M02-417, nor any indication of the support for its decision, the factors considered, who in the agency had been consulted, or how this guidance was to be communicated.

- **FDA removed key issues from the agenda of the March 2003 ACPS meeting and directed committee members not to discuss key data**: FDA first published the agenda for the March 2003 ACPS meeting on February 3, 2003. The agenda allotted time to "discuss and provide comments on levothyroxine bioequivalence." 68 FR 5297, 5298. One week before the meeting, FDA unexpectedly revised the agenda to remove reference to levothyroxine. *See* 68 FR 10254 (Mar. 4, 2003). FDA also issued a directive to committee members that the impact of Abbott's clinical data on the validity of the agency's levothyroxine BE methodology was "*not a topic for discussion at this ACPS meeting.*" Petition, Tab 17, at 656 (emphasis original). Subsequently, at the ACPS meeting, the agency announced that it had adopted a BE methodology for levothyroxine products, without any meaningful dialogue on the issue with committee members or clinicians, and in clear violation of FDA's own "Good Guidance Practice" regulations. *See* Petition at 42-45.

- **FDA belatedly acknowledged the need for a meeting on the issue of levothyroxine BE testing and solicited an agenda, but then abruptly postponed all planning for the meeting**: As discussed above, this past fall, FDA met with the Medical Societies to discuss the importance of dose precision and strict BE standards for levothyroxine products. At the meeting, FDA agreed to hold a workshop on the issues presented in Abbott's Citizen Petition. FDA has now postponed all plans for this meeting. *See* Tab D.

As these facts demonstrate, FDA has repeatedly deprived itself of relevant information from leading experts and from those who designed and carried out Study M02-417. The agency continues to lack information to engage in reasoned decision making on the core issue of levothyroxine BE methods. Moreover, the agency's inability to fulfill its fundamental obligation under the law to consider all relevant information, and to provide sponsors with an opportunity to have disputes heard before an advisory committee or panel of experts, jeopardizes the deference that FDA is ordinarily due when it engages in scientific decision making.[5]

That is, federal law accords agency action a high level of deference, especially when an agency exercises discretion in its area of scientific expertise. *See* 5 USC 706(2)(A); *see also A.L. Pharma v. Shalala*, 62 F.3d 1484, 1490 (D.C. Cir. 1995); *Schering Corp. v. FDA*, 51 F.3d 390, 399 (3d Cir. 1995); *Upjohn Co. v. Kessler*, 938 F. Supp. 439, 444 (W.D. Mich. 1996). At the same time, where courts have detected a lack of reasoned decision making, most notably a failure to consider relevant factors, they have consistently found an agency determination to be arbitrary and capricious. *See Motor Vehicle Mfrs. Ass'n of the United States v. State Farm Mut. Ins. Co.*, 463 U.S. 29, 52 (1983); *Public Citizen v. Heckler*, 653 F. Supp. 1229, 1240-41 (D.D.C. 1987). Courts also have recognized that a failure to consider relevant factors often speaks of a failure of agency process. *See, e.g., American Bioscience, Inc. v. Thompson*, 269 F.3d 1077, 1084 (D.C. Cir. 2001); *A.L. Pharma*, 62 F.3d at 1487; *United States v. Nova Scotia Food Products Corp.*, 568 F.2d 240, 251 (2d Cir. 1977); *Hanover Potato Products, Inc. v. Sullivan*, No. 1:CV-90-0746 (M.D. Pa. Aug. 3, 1990). As such, even where an agency's determination appears reasonable on its face, procedural shortcomings in its decision making process will eliminate the deference the agency usually enjoys.

---

5 On April 14, 2004, the ACPS met to discuss proposed revisions to FDA's BE methodology as applied to highly variable drugs. Abbott has been seeking advisory committee or expert consideration of the same issue for lower variability/narrow therapeutic index drugs. Clearly, FDA recognizes that its standard statistical methodology is not appropriate for all drugs, and has the time and resources to meet on such issues.

Here, the agency's refusal to have the issue of levothyroxine BE testing presented in a public forum, that allows for dialogue among the leading clinical and biopharmaceutics experts, will lead to more questions rather than definitive answers. It also will call into question the adequacy of the record in support of any agency decision with respect to levothyroxine products. Critically, the agency itself invited Abbott to submit a Citizen Petition to help "establish an administrative record on which the Agency may base any future decisions." Petition, Tab 1, at 1. Now, the agency seems intent on answering Abbott's Petition without waiting for compilation of the very administrative record it recognized it needed for sound decision making.

In this light, and in light of the pattern of conduct outlined above, FDA's apparent plan to answer Abbott's Petition *before* meeting with the leading experts is exactly wrong. Such an approach will undoubtedly call into question any and all final decisions on the subject of levothyroxine BE testing.

V. CONCLUSION

With this supplement, Abbott has added yet more evidence to the record in support of the Petition, including a declaration from a leading biostatistician. Abbott also has shown that the Hyman, Phelps Comment concedes that the agency's recommended BE methodology cannot distinguish between two levothyroxine products that differ by 12.5% or more. Finally, the agency's precipitous postponement of a meeting that FDA committed to more than seven months ago calls into question the adequacy of the administrative process being used to decide this matter.

Sincerely,

David M. Fox
Brian R. McCormick
Hogan & Hartson L.L.P.

cc: John M. Leonard, M.D.
Douglas L. Sporn
Neal B. Parker
Abbott Laboratories

Kevin M. Fain
Office of the Chief Counsel, GCF-1

FDA Docket No. 03P-0126