ABBOTT

John T. Leonard, M.D.
Vice President
Global Pharmaceutical Development
Global Pharmaceutical Research & Development

Abbott Laboratories
D R482 Bldg. AP9A
100 Abbott Park Road
Abbott Park, Illinois 60064-6120

Telephone (847) 938-5643
Fax (847) 937-0818
E-Mail: john.leonard@abbott.com

August 25, 2003

03 AUG 25 P1:51

BY HAND DELIVERY

Dockets Management Branch, HFA-305
Food and Drug Administration
5630 Fishers Lane, Room 1061
Rockville, Maryland 20852

CITIZEN PETITION

On behalf of Abbott Laboratories ("Abbott"), we submit this citizen petition pursuant to the Food, Drug, and Cosmetic Act ("FDCA") and 21 CFR 10.30 to request that the Commissioner of Food and Drugs take the actions described below to establish a clinically sensitive and scientifically valid bioequivalence ("BE") methodology for oral levothyroxine sodium drug products.

We are submitting this petition at the request of the Food and Drug Administration ("FDA"). *See* Tab 1.[1] Previously, Abbott raised its concerns about FDA's criteria for establishing the bioequivalence of levothyroxine products under the agency's formal dispute resolution process, as recommended by the Center for Drug Evaluation and Research ("CDER"). *See* Tabs 2 and 3. On May 15, 2003, CDER determined that the issues raised by Abbott were of such "significant interest" that they should be decided "in a public forum." The agency asked Abbott to submit this petition to provide the public with an opportunity to comment and to "participate in the decision-making process . . . ." Tab 1 at 1.

This petition raises issues that are central to public health. Levothyroxine is the leading treatment for hypothyroidism and the management of thyroid cancer, and is prescribed to more than 13 million Americans (nearly 1 out of

---

1 The attached documents are numbered sequentially, for ease of reference. For each reference, as appropriate, we provide both a tab number and a sequential page number (*i.e.*, Tab __ at __).

2003P-0387

CP1

every 19). The drug is safe and effective only when dosed with precision, and only when the patient's dose can be reliably maintained within a narrow range. This means that if two manufacturers' levothyroxine products differ by only a small margin, substitution of one for the other is likely to lead to therapeutic failures. In patients with coronary heart disease, in cancer patients, and in pediatric patients, a small and unexpected difference in dose presents a serious health hazard.

The ability to safely substitute among levothyroxine drug products turns on the issue of bioequivalence. The agency, however, does not have a scientifically valid BE methodology for levothyroxine products. As shown below, the agency has been advising sponsors to use a BE methodology that lacks scientific rigor. Indeed, Abbott has demonstrated with clinical data that FDA's methodology lacks the sensitivity needed to distinguish among products that release different amounts of levothyroxine. With the agency's methodology, for example, a product that delivers 112 micrograms ("mcg") of levothyroxine has a high likelihood of being declared equivalent to a product that delivers 100 mcg of levothyroxine. This difference – for products intended to be interchangeable – can cause serious harm to patients.

Neither the patients who depend on these drugs, nor the clinicians who prescribe them, can risk such uncertainty. It is incumbent upon FDA to initiate a public process – including appropriate advisory committee review – to develop an accurate, sensitive, and reproducible bioequivalence methodology for levothyroxine drug products. This petition outlines the scientific and legal basis for such a process and, critically, the need for FDA to defer any future regulatory decisions regarding the bioequivalence of levothyroxine products until a valid methodology is in place.

ACTIONS REQUESTED

On January 14, 2003, in a letter issued by the Division of Metabolic and Endocrine Drug Products (the "Division Letter"), FDA informed Abbott that the agency had adopted a "pre-dose baseline correction method" for sponsors seeking to demonstrate the bioequivalence and therapeutic equivalence ("TE") of oral levothyroxine products. Tab 4 at 151. Agency officials later announced at a March 13, 2003, meeting of the Advisory Committee for Pharmaceutical Science ("ACPS") that FDA had adopted and was recommending this methodology to sponsors. *See* Tab 5 at 198. The agency's BE methodology is not clinically supportable. Moreover, the agency adopted this method without seeking public input, without soliciting advice from an appropriate advisory committee, and without adhering to the requirements of the Administrative Procedure Act ("APA").

We therefore respectfully request that you:

(1) Immediately halt the use of the agency's BE methodology;

(2) Halt the review of any pending applications that rely on this methodology or any other methodology that has not been shown to be sufficiently sensitive to distinguish levothyroxine products that may differ by clinically relevant amounts;

(3) Refer the issue of the proper BE methodology for levothyroxine products to an appropriate advisory committee, with joint representation from the ACPS and the Endocrinologic and Metabolic Drugs Advisory Committee ("EMDAC"); and

(4) Establish, through a public process consistent with "good guidance practices" ("GGPs"), the most accurate and sensitive methodology available for demonstrating the equivalence of levothyroxine products. This methodology must include a valid method for addressing baseline levels of endogenous hormone and valid statistical criteria that take into account the narrow therapeutic range of this product.

STATEMENT OF GROUNDS

I. BACKGROUND

A. Thyroid Biology and Levothyroxine Therapy

Levothyroxine sodium is the sodium salt of the levo isomer of the thyroid hormone thyroxine. Thyroxine ("T4") is an endogenous hormone, synthesized in and released from the thyroid gland. It is a "pro-hormone" in that it is converted in the body into the more biologically potent triiodothyronine ("T3"). Both T4 and T3 affect protein, lipid, and carbohydrate metabolism, growth, and development. They stimulate the oxygen consumption of most cells of the body, resulting in increased energy expenditure and heat production. T4 and T3 also possess a cardiac stimulatory effect that may result from direct action on the heart. *See* Tab 6 at 223.

The thyroid hormone system is regulated through a tight feedback system in the hypothalamic-pituitary axis. When hormone levels are low, the hypothalamus secretes thyroid stimulating hormone-releasing hormone ("TRH"), which prompts the pituitary gland to produce thyroid-stimulating hormone ("TSH"). TSH, in turn, stimulates the thyroid gland to produce both T4 and T3. When hormone

levels are high, levothyroxine inhibits the production of TRH and TSH, decreasing hormone production in the thyroid gland. The sensitivity of this system ensures that hormone levels in healthy individuals are tightly controlled within a narrow range. *See id.* at 224.

Orally administered synthetic levothyroxine is approved for use in the treatment of hypothyroidism, as a TSH suppressant for various types of goiters, and as an adjunct to surgery and radioiodine therapy in the management of thyroid cancer. *See* Tab 7 at 261. Hypothyroidism, a condition characterized by fatigue, weight gain, memory impairment and mental retardation, brachycardia, constipation, and myalgias, among other symptoms, requires immediate treatment. Levothyroxine is prescribed for patients as young as newborns and is widely used by geriatric patients, patients with underlying coronary heart disease, and in pregnant and nursing women. In pediatric patients with congenital hypothyroidism, levothyroxine therapy must be instituted at full replacement doses to avoid retardation of intellectual and physical development. *See id.*; *see also infra* at 25-27.

As outlined by FDA, levothyroxine must be precisely and consistently dosed for it to be safe and effective:

> If a drug product of lesser potency or bioavailability is substituted in the regimen of a patient who has been controlled on one product, a suboptimal response and hypothyroidism could result. Conversely, substitution of a drug product of greater potency or bioavailability could result in toxic manifestations of hyperthyroidism such as cardiac pain, palpitations, or cardiac arrhythmias. In patients with coronary heart disease, *even a small increase in the dose of levothyroxine sodium may be hazardous.*

62 FR 43535, 43536 (Aug. 14, 1997) (emphasis added); *see* Tab 8 at 327 (FDA discussing the numerous "serious adverse consequences if the dose [of levothyroxine] is not specifically titrated to the needs of the individual patient" and if the dose is not precisely maintained throughout the course of treatment).

Based on the need for precise dosing, FDA has expressly recognized the clinical significance of dosing increments as low as 12 mcg for levothyroxine sodium products. According to the agency, "[l]evothyroxine sodium products are marketed in multiple dosage strengths [*i.e.*, 25, 50, 75, 88, 100, 112, 125, 137, 150, 175, 200, and 300 mcg], that may vary by only 12 micrograms, thus permitting careful titration of dose. Because of levothyroxine sodium's narrow therapeutic index, it is particularly

important that the amount of available active drug be consistent for a given tablet strength." 62 FR at 43538.[2]

Indeed, the labeling for levothyroxine products recommends dosing adjustments of 12.5 to 25 mcg for elderly patients with underlying cardiac disease and patients with severe hypothyroidism. *See* Tab 7 at 274 ("The levothyroxine sodium dose is generally adjusted in 12.5-25 mcg increments until the patient with primary hypothyroidism is clinically euthyroid and the serum TSH has normalized."); *see also* Tab 8 at 327-28 ("[A] 25 mcg dosage strength . . . is essential for proper labeling of the product for safe and effective use given that in certain clinical situations, levothyroxine sodium dosing is initiated at 12.5-25 mcg/day and increased in 12.5-25 mcg dosing increments.").

B. The Approval of Oral Levothyroxine Products

For decades, dating back to the 1950s, oral levothyroxine products generally were distributed without premarket approval by FDA. In August 1997, FDA published a notice in the *Federal Register* stating that the agency would, going forward, regulate all oral levothyroxine products as "new drugs" for which premarket approval is required under section 505 of the FDCA. *See* 62 FR 43535.

FDA's basis for taking this action was the concern that levothyroxine products lacked consistent potency and bioavailability ("BA"). According to the agency, patients were being exposed to products of varying potency, and this variation was the cause of numerous therapeutic failures. *See id.* at 43536; Tab 5 at 194. In support of its decision, FDA cited 58 reports of adverse drug events associated with levothyroxine products, nearly half of which occurred following the refill of a prescription. *See* Tab 5 at 194.

To remedy this problem, the agency asserted regulatory control over levothyroxine products and called for the submission of new drug applications ("NDAs") and abbreviated new drug applications ("ANDAs") within three years (later extended to four years). With greater control over the formulation and manufacturing of levothyroxine products, the agency could ensure that – from refill-to-refill – patients would continue to receive precisely the same dose to which they had been titrated.

---

[2] A "narrow therapeutic index" drug is one that is subject to therapeutic drug concentration or pharmacodynamic monitoring, and/or where product labeling indicates a narrow therapeutic index designation. *See* Guidance for Industry: *Bioavailability and Bioequivalence Studies for Orally Administered Drug Products – General Considerations* (Mar. 2003) at 20 ("General BA/BE Guidance").

Since August 1997, seven sponsors have gained approval of brand-name levothyroxine products. Abbott's product, Synthroid®, gained approval on July 24, 2002.[3] *See* Tab 7 at 278. These products are listed in FDA's publication, *Approved Drug Products with Therapeutic Equivalence Evaluations* (the "*Orange Book*"), and all have been designated by FDA as "reference standards" for the approval of generic levothyroxine products.[4] The agency also has approved one generic levothyroxine product with an "AB" therapeutic equivalence rating to a brand-name product (Jerome Stevens's Unithroid).[5]

C. FDA's Levothyroxine Bioavailability Guidance

Following the August 1997 notice, the agency issued a series of guidance documents to assist sponsors in gaining new drug approval for levothyroxine products.[6] One such document was a guidance on acceptable methods for demonstrating the bioavailability of oral levothyroxine products. *See* Guidance for Industry: *Levothyroxine Sodium Tablets – In Vivo Pharmacokinetic and Bioavailability Studies and In Vitro Dissolution Testing* (Feb. 2001) (the "BA Guidance" or the "guidance") (attached at Tab 10). With at least 37 entities then marketing levothyroxine products, FDA issued these documents to provide consistent advice to the numerous manufacturers expected to submit applications. *See* Tab 5 at 193-94.

---

3 Abbott's predecessor, Knoll Pharmaceuticals ("Knoll"), challenged the agency's August 1997 determination that Synthroid® is a "new drug." In a citizen petition dated December 15, 1997, Knoll argued that Synthroid® meets the "general recognition" standard under section 201(p) of the FDCA and, therefore, does not require approval under an NDA. *See* FDA Docket No. 97N-0314. On April 26, 2001, the agency denied the petition. *See* Tab 9 at 335. Rather than challenge that denial, Abbott agreed to submit an NDA in support of Synthroid®.

4 There are two citizen petitions and a petition for stay pending before FDA that challenge the legality of FDA's decision to designate each approved levothyroxine product as a reference listed drug. *See* FDA Docket Nos. 03P-0097 and 03P-0210.

5 An "AB" therapeutic equivalence rating represents FDA's determination that two or more pharmaceutically equivalent products "can be substituted with the full expectation that the substituted product will produce the same clinical effect and safety profile as the prescribed product." *Orange Book* at viii.

6 *See, e.g.,* Guidance for Industry: *Levothyroxine Sodium Products Enforcement of August 14, 2001 Compliance Date and Submission of New Applications* (July 2001); Guidance for Industry: *Levothyroxine Sodium Questions and Answers* (Feb. 2001).

The BA Guidance was "intended to assist sponsors of new drug applications (NDAs) for levothyroxine sodium products who wish to conduct in vivo pharmacokinetic and bioavailability studies and in vitro dissolution testing for their products." Tab 10 at 353.[7] In the guidance, FDA recognized that the presence in the body of baseline levels of endogenous or naturally-occurring thyroid hormone may represent a significant confounding factor in conducting levothyroxine BA studies. As the agency stated, "[i]t is a challenge to determine the bioavailability of levothyroxine sodium products because levothyroxine is naturally present in minute quantities in the blood, with the total levels reaching 5.0-12.0 [mcg/deciliter ("dl")] and free (or unbound) levels reaching 0.8-2.7 [nanograms]/dl in a healthy adult." *Id.* at 354.

The guidance, however, does not recommend that sponsors seek to correct for this baseline. The baseline levels, according to the guidance, are too "unpredictable during the course of study" to allow for correction. *Id.* at 356. Instead, the guidance recommends the use of "several times the normal dose" of levothyroxine, to raise the level of the drug sufficiently above baseline to allow for valid measurement. *Id.* at 354. The guidance recommends, but does not require, a 600 mcg test dose, *i.e.*, "a multiple of the highest tablet strength . . . ." *Id.* at 355.

D. Abbott's Analysis of the Impact of Baseline T4

1. Abbott's Concerns Regarding the Agency's Recommended Bioavailability Methodology

To obtain approval of the NDA for Synthroid®, Abbott followed the approach recommended in the BA Guidance for demonstrating the BA and pharmacokinetics of Synthroid®. However, in the course of developing its application, Abbott began to focus on the agency's decision *not* to incorporate a baseline correction method into its guidance.

As noted above, the BA Guidance recognizes that naturally-occurring levothyroxine is present in the blood of a healthy adult in amounts ranging from 5.0 to

---

[7] FDA has made inconsistent statements with respect to the applicability of this guidance to sponsors seeking approval of generic levothyroxine products on the basis of a showing of bioequivalence. In approving the first generic product in this class, the agency referred to the BA Guidance as the appropriate source for information on BE criteria. *See* Tab 11 at 477. In contrast, at a recent ACPS meeting, the Director of the Division of Bioequivalence in the Office of Generic Drugs emphasized that the guidance "refers only to the bioavailability of [levothyroxine] products. It does not address the bioequivalence [of levothyroxine products]." Tab 5 at 190; *see also id.* at 194 ("This guidance was intended to address issues of bioavailability . . . and was never intended to be used on its own for the purposes of bioequivalence.").

12.0 mcg/dl. *See* Tab 10 at 354. In Abbott's BA studies, conducted with 600 mcg doses, the maximum observed levothyroxine concentration (*i.e.,* endogenous + exogenous) reached approximately 14 mcg/dl. *See* Tab 12 at 504. Thus, in the case of a test subject with a mean endogenous T4 level of 8 mcg/dl, the baseline would account for nearly 60 percent of the total peak level of levothyroxine.

This, Abbott recognized, could lead to significant error when comparing the bioavailability of two levothyroxine products. Consider, for example, a study in which the mean baseline endogenous T4 level is 8 mcg/dl during a 24-hour cycle. One product yields total mean T4 concentrations (endogenous + exogenous) of 14 mcg/dl, while the other product yields total mean concentrations of 15 mcg/dl. Without accounting for baseline, the two products would show a difference of only 7 percent (*i.e.*, 15 *versus* 14 mcg/dl). With an adjustment for baseline, however, the two products are shown to differ by 17 percent (*i.e.*, 7 *versus* 6 mcg/dl). Because such a large proportion of the observed levels is common to both concentrations, in the form of endogenous T4, failure to account for baseline can lead to an erroneous declaration of bioequivalence.

In the course of investigating the impact of baseline T4, Abbott also became concerned that levothyroxine BA assessments may be complicated by other factors related to the effect that exogenous doses can have on baseline levels. As noted above, low circulating hormone levels lead to increased production of T4 and T3, while relatively higher levels (*e.g.*, after an exogenous dose) result in reduced hormone production. Diurnal hormone cycles also affect endogenous levels, which in turn may impact the body's natural feedback mechanism. Finally, Abbott included in its investigation an analysis of the "carryover" effects of large doses of T4 in crossover BA studies. These effects can vary, depending on the test dose of levothyroxine and the length of a study's washout period, and can further bias the final results of a BE study.

2. The Abbott Simulation Study

In November 2001, Abbott decided to initiate a series of studies to evaluate the true impact in BE studies of baseline levothyroxine and the additional factors described above. The first study involved a stochastic simulation, based on existing data, to evaluate the impact of baseline levothyroxine levels. Abbott based the simulation on data generated in support of the Synthroid® NDA, including the dosage form proportionality studies conducted by Abbott according to the BA Guidance and FDA's general BE criteria.[8] *See* Tab 13 at 515.

---

[8] Under FDA's general BE criteria, BE data are analyzed based on a 90 percent confidence interval and 80 to 125 percent acceptance criteria. *See* Guidance for Industry: *Statistical Approaches to Establishing Bioequivalence* (Jan. 2001).

The simulation, conducted for Abbott by Thomas Ludden, Ph.D., of GloboMax LLC, compared Abbott's uncorrected BA data to baseline corrected data. In carrying out the simulation, Dr. Ludden tested two different assumptions with respect to the impact that orally administered levothyroxine has in the body. These assumptions were intended to represent the outer limits of the potential impact of orally administered levothyroxine on endogenous hormone levels. *See id.*

First, Dr. Ludden prepared an analysis based on the assumption that exogenous levothyroxine has no suppressive effect on the body's natural production of levothyroxine. Under this assumption, baseline (endogenous) levels measured prior to dosing are assumed to be the same as baseline levels present after dosing. *See id.* Second, Dr. Ludden prepared an analysis based on the assumption that exogenous levothyroxine has a completely suppressive effect on the body's natural production of levothyroxine, and that any endogenous levothyroxine that is present prior to dosing declines according to its elimination half-life. *See id.* Using these two approaches, Dr. Ludden estimated by stochastic simulation the probability of two products that differ by up to 50 percent being declared equivalent under FDA's standard BE methodology.[9]

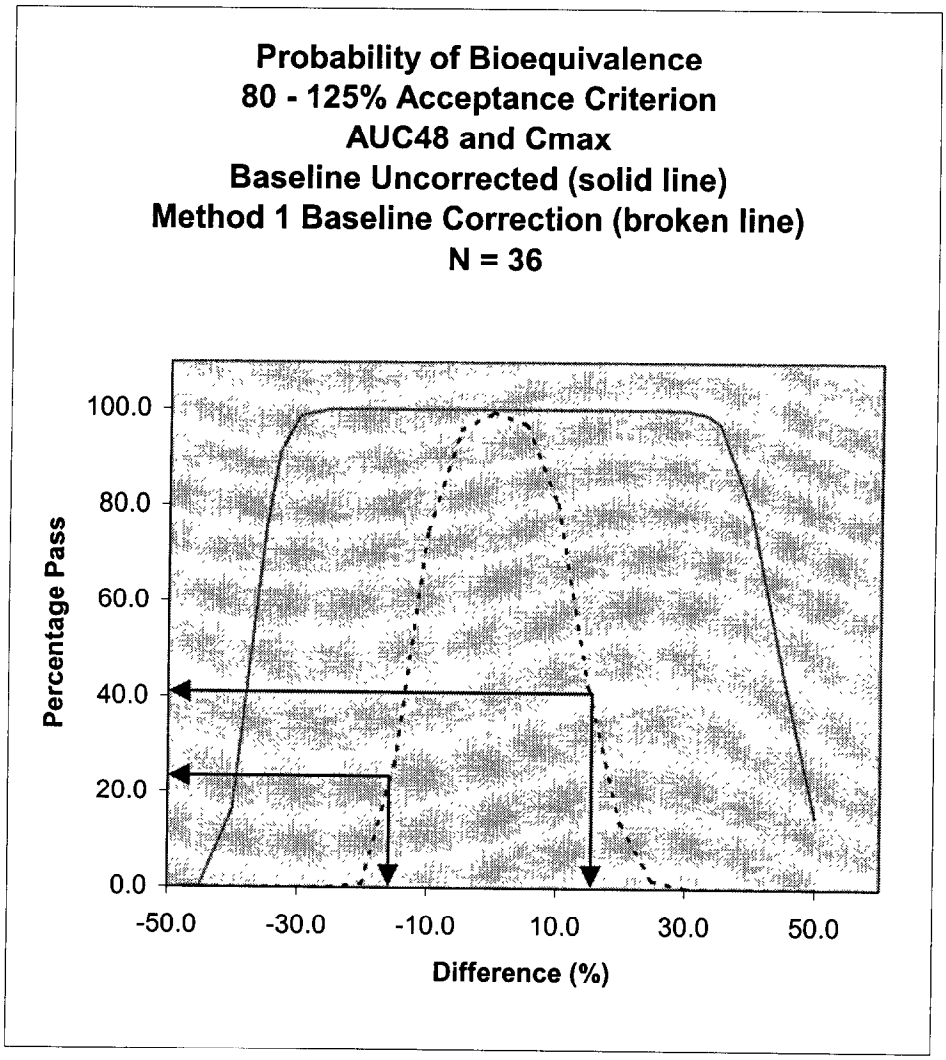
Operating under the first assumption, Dr. Ludden corrected the post-dose BA data by subtracting out the average of three pre-dose measurements of baseline levothyroxine. His analysis showed that, with this correction method, a test product that delivers 15 percent less (or more) levothyroxine than a reference product would have a 26 (or 42) percent chance of being declared bioequivalent to the reference product (by both C_{max} and AUC)[10] in a 36-subject study, based on 80 to 125 percent acceptance criteria. *See id.* at 612.

---

9 Similar studies have been used in the past by FDA scientists to evaluate the performance of BE assessments. *See* Tab 14 (attaching F.Y. Bois, et al., *Bioequivalence: Performance of Several Measures of Rate of Absorption*, Pharm. Res. 11: 966-74 (1994); F.Y. Bois, et al., *Bioequivalence: Performance of Several Measures of Extent of Absorption*, Pharm. Res. 11: 715-22 (1994)).

10 C_{max} measures the peak concentration of a drug substance in the blood. AUC measures total systemic drug exposure over a given period. For example, AUC_{48} measures exposure for 48 hours from the initial dosing. Both C_{max} and AUC are standard measurements used to gauge the release and absorption of an active drug substance in the body. *See* Guidance for Industry: *Statistical Approaches to Establishing Bioequivalence* (Jan. 2001); *see also infra* at 20.

Dr. Ludden's findings are illustrated in the graph below, which shows the probability of declaring two products bioequivalent (by both C_{max} and AUC) as a function of the true percentage difference between the products.



See id.[11]

Operating under the second assumption, Dr. Ludden corrected the post-dose BA data by subtracting the remaining baseline levothyroxine as it is eliminated

---

11 This graph presents the joint probability of a finding of bioequivalence, based on both C_{max} and AUC_{48}. Graphs depicting the probabilities for C_{max} and AUC_{48} individually are attached. *See* Tab 13 at 533.

from the body. With this correction method, his analysis showed that a product that delivers 15 percent less (or more) levothyroxine than a reference product would have a 33 (or 57) percent chance of being declared bioequivalent to the reference product (by both C_{max} and AUC) in a 36-subject study. *See id.* at 615.

Abbott submitted Dr. Ludden's analysis to FDA on May 8, 2002, as "Simulation Study to Assess Alternative Bioavailability Calculations, Study Designs and Acceptance Criteria for Determining the Bioequivalence of Levothyroxine Sodium Tablets." Tab 15 at 638. Abbott also requested a meeting with the relevant FDA clinical and biopharmaceutics experts, to discuss the simulation study and Abbott's plan for a clinical study to test its results. *See id.* at 634-35.

3. The Abbott Pharmacokinetic Study

In light of the significant findings of the simulation study, Abbott next sought to test the importance of baseline correction in a clinical setting.

In March 2002, Abbott initiated a clinical pharmacokinetic ("PK") study ("Study M02-417") to test whether levothyroxine products that actually differ by clinically relevant amounts could indeed pass as bioequivalent if the investigator failed to account for baseline levothyroxine. Abbott based the design of Study M02-417 on the agency's BA Guidance using, as FDA directed, dosing levels "several times the normal dose" to raise the level of endogenous hormone high above baseline. Tab 10 at 354. Thus, the first arm of the study used a 600 mcg dose, the second arm used a 450 mcg dose, and the third arm used a 400 mcg dose. Blood samples were collected as per the guidance, with additional samples taken to more fully assess endogenous levothyroxine levels. Samples were collected for 24 hours prior to, and out to 96 hours after, dosing. Finally, per the guidance, the relevant PK measures (C_{max}, T_{max}, and AUC_{48}, AUC_{72} and AUC_{96}) were analyzed without baseline correction.[12] *See* Tab 12 at 501.

As predicted by Dr. Ludden's simulation study, each PK measure in Study M02-417 – when analyzed without baseline correction – was consistent with a finding of bioequivalence. Test and reference doses that differed by 12.5 percent, 25

---

12 Abbott took numerous steps to control for all possible study variables. For example, all doses of the study drug came from the same production lot, and were aged to the same date. Content uniformity data for the lot demonstrated a narrow range of individual tablet potency values at the time of release. In addition, all serum analytes were assayed in batches, such that all samples from any one test subject were measured in a single batch, reducing the possibility that any individual's measurements were the result of inter-batch variability. *See* Tab 12 at 502.

percent, and 33 percent were found to be bioequivalent.[13] *See id.* at 504. Abbott then analyzed the data using each of three baseline correction methods, to determine whether the BE methodology could be refined to distinguish inequivalent products. A description of the methods evaluated by Abbott follows:

> **Method 1**: The pre-dose baseline value on the day of dosing was subtracted from each post-dose concentration. The pre-dose baseline value was calculated as the average of three concentrations (at 0.5, 0.25, and 0 hours) taken prior to dosing in each period. (This method assumes no suppression of endogenous T4 production, and is the same as that used in Abbott's simulation study.)
>
> **Method 2**: For each time of post-dose sampling, the observed concentration was corrected assuming that the endogenous T4 baseline at 0 hours declined according to a half-life of 7 days. (This method assumes complete suppression of endogenous T4 production, and also was evaluated in Abbott's simulation.)
>
> **Method 3**: The T4 concentration for each time of post-dose sampling was corrected by the concentration observed at the same time of day during the 24 hours preceding the dose. (This method allows for a diurnal hormone cycle, but assumes that it is not changed by the administration of the levothyroxine dose.)

See id. at 503.

Each baseline correction method reduced the likelihood that two products differing by 25 to 33 percent would be found BE. However, none of the three methods, used in conjunction with FDA's standard BE methodology, was sufficiently sensitive to distinguish the doses that differed by 12.5 percent. For all but one PK measure (under Correction Method 3), the 450 and 400 mcg doses would easily be declared bioequivalent under FDA's standard analysis. *See id.* at 506-07.

---

[13] For C_{max} and each measure of AUC, the 90 percent confidence intervals for the ratios of central values were fully contained within FDA's standard acceptance criteria, 80 to 125 percent. The analysis was performed on logarithmically transformed data, using the appropriate analysis of variance model. For example, for doses that differed by 33 percent (400 mcg *versus* 600 mcg), the confidence intervals for C_{max} and AUC_{48} were 0.88 to 0.96 and 0.90 to 0.96, respectively. For doses that differed by 12.5 percent (450 mcg *versus* 400 mcg), the confidence intervals for C_{max} and AUC_{48} were 0.97 to 1.05 and 1.00 to 1.06, respectively. *See id.* at 504.

The results, using Correction Method 1, were as follows:[14]

Bioequivalence and Relative Bioavailability for T_4 (Correction Method 1)

Regimens Test *vs.* Reference	Pharmacokinetic Parameter	Central Value* Test	Central Value* Reference	Relative Bioavailability Point Estimate+	Relative Bioavailability 90% Confidence Interval
450 mcg	C_{max}	5.4	6.9	0.783	0.727 – 0.844
vs.	AUC_{48}	119.7	167.3	0.715	0.658 – 0.778
600 mcg	AUC_{72}	151.4	215.7	0.702	0.636 – 0.774
	AUC_{96}	170.2	250.2	0.680	0.602 – 0.768
400 mcg	C_{max}	5.6	6.9	0.803	0.745 – 0.865
vs.	AUC_{48}	118.9	167.3	0.711	0.653 – 0.773
600 mcg	AUC_{72}	144.9	215.7	0.672	0.609 – 0.741
	AUC_{96}	165.1	250.2	0.660	0.584 – 0.746
450 mcg	C_{max}	5.4	5.6	0.975	0.906 – 1.049
vs.	AUC_{48}	119.7	118.9	1.007	0.926 – 1.094
400 mcg	AUC_{72}	151.4	144.9	1.044	0.948 – 1.150
	AUC_{96}	170.2	165.1	1.031	0.914 – 1.163

* Antilogarithm of the least squares means for logarithms.
\+ Antilogarithm of the difference (test minus reference) of the least squares means for logarithms.

Thus, the 450 and 400 mcg doses "passed" bioequivalence by a comfortable margin. Although Dr. Ludden's original simulation study did not estimate the probability that two products differing by 12.5 percent would be found bioequivalent, Dr. Ludden subsequently re-ran the simulation to examine the probability of such a finding. According to his analysis, a test product that delivers 12.5 percent less (or more) levothyroxine than a reference product would have a 52 (or 62) percent chance of being declared BE to the reference product (by both AUC and C_{max}) in a 36-subject study, under Correction Method 1, using 80 to 125 percent acceptance criteria. *See supra* at 29. Indeed, given the margins by which the 450 and 400 mcg doses were shown to be BE, and given the results of Dr. Ludden's simulation, it is likely that this BE methodology also would fail to distinguish doses that differ by more than 12.5 percent. *See* Tab 12 at 506.

---

[14] Summary tables for Correction Methods 2 and 3 are attached. *See id.* at 506-07.

E. FDA's Response to Abbott's Submissions

1. FDA's Decision to Adopt a Baseline Correction Method and Abbott's Appeal of that Decision

Abbott submitted its simulation study to FDA, along with a request for a meeting, in early May 2002. *See* Tab 15. On May 20, 2002, CDER's Division of Metabolic and Endocrine Drug Products (the "Division") denied the meeting request and stated that it would reconsider the request after Abbott submitted a final clinical study report. *See* Tab 2 at 36. On October 10, 2002, the company submitted its final clinical study report for Study M02-417 and renewed its request for a meeting. *See id.* at 40.

On January 14, 2003, the Division issued a letter to Abbott stating that there was no need for a meeting and that FDA had decided the matter. *See* Tab 4. The agency stated that it had adopted "a three pre-dose baseline subtraction method to evaluate total thyroxine" when considering products for "AB" therapeutic equivalence ratings, and that it would recommend the method to all levothyroxine sponsors. *Id.* at 151. The letter provided no explanation in support of the decision and no indication as to who had been consulted, what factors were considered, or how this guidance was being communicated. Nor did the letter address the data from Study M02-417, showing that such a correction method cannot distinguish doses that differ by 12.5 percent or more.

Per the agency's recommendation that Abbott initiate "formal dispute resolution" should it disagree, the company promptly appealed the decision. Tab 2 (the "February 12 FDR Submission").[15] In addition to presenting the data and analysis from Study M02-417, Abbott demonstrated that, as a clinical matter, failure to distinguish between doses that differ by 12.5 percent (or less) can have serious adverse health consequences for thyroid patients. *See id.* at 15-17. Abbott requested a review of the decision and a joint advisory committee meeting – with the relevant clinical and biopharmaceutics experts – to review the issue of the proper BE methodology for levothyroxine products. *See id.* at 19.

On February 13, 2003, Abbott met with members of CDER's Office of Pharmaceutical Science and Office of Generic Drugs to prepare for an upcoming meeting of the ACPS on March 12-13, 2003. A portion of the advisory committee

---

15 *See* Guidance for Industry: *Formal Dispute Resolution: Appeals Above the Division Level* (Feb. 2000) ("FDR Guidance").

meeting was scheduled to cover endogenous drugs, and Abbott therefore planned to present the results of Study M02-417 to the ACPS.[16] At the February 13 meeting, Abbott provided CDER with a proposed slide deck and summarized the presentation it expected to make before the advisory committee. FDA asked several questions about the quality of the drug used in Study M02-417 and the statistical and clinical significance of the findings, and requested several revisions to Abbott's slides. The agency, however, offered no comments or criticisms on the design of Study M02-417. For example, the agency gave no indication that it disagreed with Abbott's findings with respect to the 450 and 400 mcg doses. *See infra* at 31-35. The agency also did not provide Abbott with its own slide deck or an outline of its proposed remarks to the committee.

2. The Response to Abbott's Appeal

On March 7, 2003, CDER's Office of Drug Evaluation II ("ODE II") responded to Abbott's request for dispute resolution. *See* Tab 16 (the "ODE II Letter"). The ODE II Letter affirmed the agency's decision to adopt a three pre-dose baseline correction method and denied Abbott's request for a joint advisory committee meeting. According to the letter, Abbott's data – showing that FDA's BE methodology cannot distinguish a 450 mcg dose of levothyroxine from a 400 mcg dose – are invalid. These doses are, according to the letter, "well below the 600 mcg dose" recommended in the BA Guidance. *Id.* at 645. According to FDA, at 450 and 400 mcg, the baseline "noise" drowns out the dose "signal." *Id.* The letter, however, cites no data, literature, or analysis to support this assertion.

With respect to Abbott's request for a joint clinical/biopharmaceutics advisory committee meeting, CDER denied the request. According to the ODE II Letter, the clinical issues are well understood: FDA "believe[s] the clinical importance of levothyroxine and having the correct dosage is very clear to the Agency's own medical experts . . . ." *Id.* As to the biopharmaceutics issues, CDER stated that these would be sufficiently covered during the upcoming March 13, 2003, ACPS meeting. *See id.* at 646.

The ODE II Letter was dated March 7, 2003. The letter, however, was sent by ordinary mail and did not arrive at Abbott's headquarters in time for the company to review it before the March 13 ACPS meeting. FDA did not take steps to

---

16 The ACPS meeting was scheduled prior to Abbott's initiation of formal dispute resolution. As planned and as carried out, the ACPS meeting was categorically different from the meeting sought in Abbott's request for formal dispute resolution. *See infra* at 16-17, 39.

ensure that Abbott received the letter promptly (*e.g.*, by sending a courtesy copy by facsimile, overnight delivery, or e-mail), despite Abbott's request, made in its February 12 FDR Submission, for an explanation of FDA's reasoning in order "to make for a more productive advisory committee meeting . . . ." Tab 2 at 4.

Finally, as with the Division Letter, the ODE II Letter closed with an invitation to appeal the decision to the next supervisory level in CDER, which Abbott promptly did. *See* Tab 3 (the "April 14 FDR Submission").

3. The March 13, 2003, Advisory Committee Meeting

On February 3, 2003, FDA published its agenda for the next regularly scheduled meeting of the Advisory Committee for Pharmaceutical Science, to be held on March 12 and 13, 2003. The agenda included the following item: "[D]iscuss and provide comments on levothyroxine bioequivalence." 68 FR 5297, 5298 (Feb. 3, 2003). In the following weeks, however, FDA took several steps to reduce the focus on levothyroxine.

Just eight days before the meeting, the agency re-published the agenda, removing the reference to levothyroxine products, so that it read, "discuss and provide comments on bioequivalence/bioavailability of endogenous drugs." 68 FR 10254 (March 4, 2003). Several days after that, FDA informed the ACPS members that Abbott's clinical data, and the validity of the agency's levothyroxine BE methodology, were *not* to be discussed at the meeting:

> A few months ago Abbott Labs provided the agency data from a study related to the FDA guidance "Levothyroxine Sodium Tablets – In Vivo Pharmacokinetic and Bioavailability Studies and In Vitro Dissolution Testing." This study illustrates several aspects that need to be considered with respect to study design and data analysis of endogenous drugs. We have, therefore, invited them to share this information with you. Abbott has raised with FDA some issues related to the impact of their study results on the bioequivalence assessment of levothyroxine. *This is not a topic for discussion at this ACPS meeting.* During the open public session several speakers have requested time to express their opinions on the issue of bioequivalence of levothyroxine products. Again, these do not directly apply to this discussion. The FDA welcomes these opinions and will collect these for consideration in an appropriate manner.

Tab 17 at 656 (emphasis in original). Finally, in briefing materials posted in advance of the meeting, and during the meeting itself, the agency made clear that the issue of

levothyroxine was to be considered "an awareness topic" only and not a substantive topic for advisory committee consideration. *Id.* at 653, 656; Tab 5 at 190.

Nevertheless, the agency used the meeting to announce – for the first time to a public audience – that FDA had adopted a BE methodology for oral levothyroxine products.[17] FDA's "preferred method" of baseline correction was described as subtracting the mean of three pre-dose levothyroxine concentrations from all subsequent post-dose concentrations. Tab 5 at 199. This method is the same as Correction Method 1 in Abbott's Study M02-417, discussed earlier. *See supra* at 12. The agency provided neither Abbott nor the ACPS members any advance notice, briefing materials, or proposed slides on its presentation.

With regard to Study M02-417, a senior CDER official stated at the meeting that the study was "very useful when the FDA decided to adopt a baseline correction method," but that the balance of the study was essentially invalid. Tab 5 at 198. In particular, the agency argued that Abbott's use of 450 and 400 mcg doses yielded blood levels that were too close to baseline, and that baseline interference prevented an "accurate evaluation of the true differences that exist between the two doses . . . ." *Id.* FDA presented no data in support of this hypothesis.

At the meeting, practicing clinicians testified about the unique sensitivity of thyroid patients to small changes in levothyroxine dose and about the need for a clinically sensitive BE methodology. *See id.* at 178-89.[18] However, the ACPS is not composed of clinical experts in endocrinology. FDA also failed to provide the committee, in advance of the meeting, with the agency's analysis of the levothyroxine BE issue. Finally, the agency chose not to present any questions to the clinicians who testified and did not solicit any recommendations from the committee. As a result, there was no meaningful dialogue with the committee or with the clinicians at the meeting.

In short, without advance notice of FDA's arguments, neither Abbott nor the advisory committee was able to engage the agency in any kind of meaningful dialogue. The scientific exchange that CDER described in the ODE II Letter (*see* Tab 16 at 645-46) failed to take place.

---

17 The transcript of the ACPS meeting and copies of the relevant slides presented by FDA are attached at Tab 5. These materials and all other slide decks and handouts used at the meeting are available at www.fda.gov/ohrms/dockets/ac/cder03.html and are incorporated herein.

18 The clinicians' views are discussed in greater depth below. *See infra* at 25-27.

4. Abbott's Second Request for Dispute Resolution and FDA's Request for a Public Process

Per the recommendation in the ODE II Letter, Abbott continued its appeal of FDA's decision to adopt a pre-dose baseline correction method. On April 14, 2003, Abbott submitted to the next supervisory level within CDER, the Office of New Drugs, an extensive analysis of the scientific and procedural issues associated with the agency's decision. *See* Tab 3.

Abbott's argument again centered on the agency's failure to address the clinical issues associated with small differences in levothyroxine doses. In particular, Abbott explained in the April 14 FDR Submission that the ODE II Letter omitted any showing that the agency's BE methodology can distinguish products that differ by clinically relevant amounts. The April 14 FDR Submission also focused on the procedural defects associated with the agency's decision to adopt a BE methodology for this class of products without following GGPs and without advisory committee review. *See id.* at 70-72.

On May 15, 2003, CDER's Office of Regulatory Policy issued a letter to Abbott that addressed, in part, the need for a public process. *See* Tab 1. The letter explained that while FDA had initially invited Abbott to pursue formal dispute resolution, the agency now recognized that the issues raised by Abbott are of "significant interest to other manufacturers of levothyroxine sodium, including generic drug applicants." *Id.* at 1. According to the agency,

> [W]e believe it is most appropriate to consider the issues raised in your request in a public manner. This approach will allow others the opportunity to comment and participate in the decision-making process, will provide Abbott the opportunity to comment publicly on the views and opinions of others, and will establish an administrative record on which the Agency may base any future decisions.

Id. Thus, the agency asked Abbott to forego its formal dispute resolution in favor of submitting a public citizen petition (or commenting on a related petition submitted by Jones Pharma, Inc.). *See* Tab 18. Abbott promptly notified the agency that it would submit an original petition. *See* Tab 19.

II. STATUTORY AND REGULATORY FRAMEWORK

Under the FDCA, any person seeking to market a generic drug under an ANDA must show, among other things, that the proposed generic is bioequivalent to a

reference drug. *See* 21 USC 355(j)(2)(A)(iv). By statute, a generic drug is bioequivalent to the reference, or "listed," drug if:

> [T]he rate and extent of absorption of the [proposed] drug do not show a significant difference from the rate and extent of absorption of the listed drug when administered at the same molar dose of the therapeutic ingredient under similar experimental conditions in either a single dose or multiple doses . . . .

Id. at 355(j)(8)(B)(i); *see* 21 CFR 320.1.

Drug products that are determined to be bioequivalent and pharmaceutically equivalent (*i.e.*, they contain the identical active ingredient in the identical amount and dosage form, 21 CFR 320.1(c)) are eligible to be classified by FDA as "therapeutically equivalent." Therapeutically equivalent products, according to FDA, "can be substituted with the full expectation that the substituted product will produce the same clinical effect and safety profile as the prescribed product." *Orange Book* at viii.[19]

Thus, the methodology that an applicant uses to demonstrate bioequivalence is critical. It must be sensitive enough to detect a "significant difference" between the test and the reference products. A methodology that cannot detect a significant difference in the rate and extent of absorption between the test and the reference products necessarily fails to meet the statutory BE standard. *See* 21 USC 355(j)(8)(B)(i).

By regulation, FDA has set forth the types of evidence that may be used to demonstrate bioequivalence. *See* 21 CFR 320.24. More detailed recommendations on how to demonstrate bioequivalence for certain categories of products, such as oral tablets and capsules, are outlined in agency guidance. *See, e.g.*, General BA/BE Guidance. Also, for specific classes of drug products that pose challenges with respect to bioavailability and bioequivalence, the agency may issue product-specific guidance. *See, e.g.*, Draft Guidance for Industry: *Potassium Chloride Modified-Release Tablets*

---

19 Sponsors also may seek to submit NDAs with bioequivalence data under section 505(b)(2) of the FDCA. Abbott does not believe, however, that products approved under section 505(b)(2) are eligible to be classified as therapeutically equivalent by FDA. *See* Abbott Comments to FDA Docket No. 01P-0323 (July 10, 2002). Moreover, FDA has indicated that it will refuse to file 505(b)(2) applications for levothyroxine sodium products that are eligible for approval under section 505(j). *See* 21 CFR 314.101(d)(9); Guidance for Industry: *Levothyroxine Sodium Products Enforcement of August 14, 2001 Compliance Date and Submission of New Applications* (July 2001) at 4. For these reasons, we refer in this petition only to applications under section 505(j).

and Capsules: In Vivo *Bioequivalence and* In Vitro *Dissolution Testing* (Aug. 2002); Guidance for Industry: *Clozapine Tablets* In Vivo *Bioequivalence and* In Vitro *Dissolution Testing* (Nov. 1996); Guidance for Industry: *Phenytoin/Phenytoin Sodium Capsules, Tablets and Suspension* In Vivo *Bioequivalence and* In Vitro *Dissolution Testing* (Mar. 1994).

At a minimum, a bioequivalence study must include three elements. First, it must include appropriate measures by which to compare the release and absorption of the active drug substance from each product. For oral tablets, the ratios of the peak concentrations in blood ("C_{max}") and systemic exposures ("AUC") are generally considered to be appropriate measures. *See* Guidance for Industry: *Statistical Approaches to Establishing Bioequivalence* (Jan. 2001) at 2. Second, the methodology must include an acceptable confidence interval to assess the statistical significance of these measures. Generally, the agency accepts a 90 percent confidence interval for each log transformed pharmacokinetic measure (*i.e.,* C_{max} and AUC).

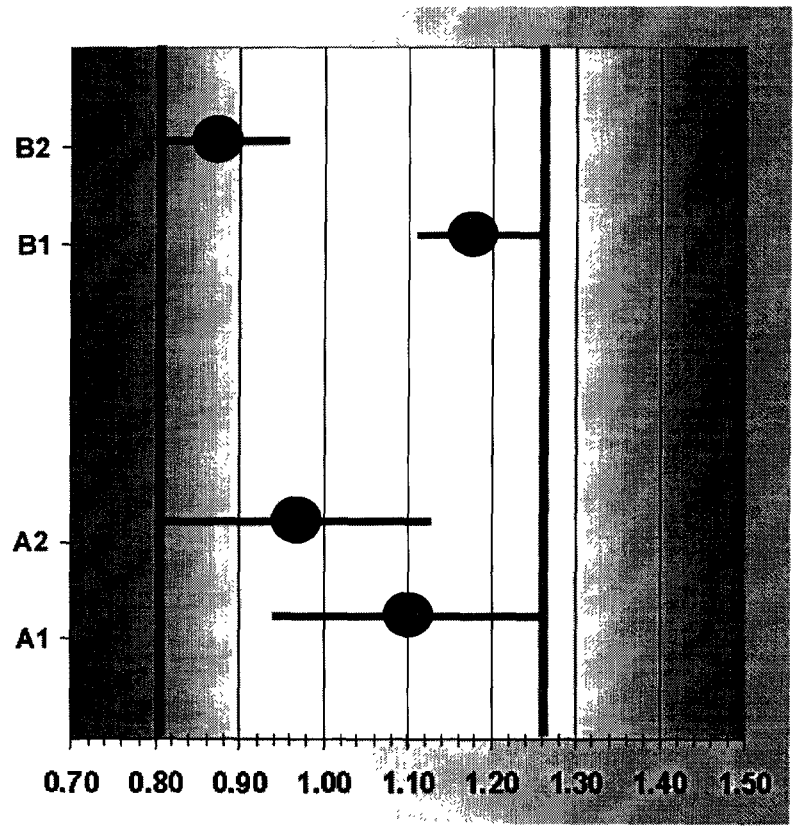
Third, the methodology must include a bioequivalence range or "acceptance criteria." As noted above, the test and reference products can differ in systemic absorption, so long as the difference is not "significant." 21 USC 355(j)(8)(B)(i). The agency generally accepts an 80 percent lower limit and a 125 percent upper limit, within which the entire 90 percent confidence interval must fall. *See* Guidance for Industry: *Statistical Approaches to Establishing Bioequivalence* (Jan. 2001) at 2. The agency, however, has recognized that tighter limits may be appropriate where the real differences permitted under this methodology would be clinically significant. *See, e.g.,* General BA/BE Guidance at 20; *see also infra* at 36.

As FDA recognizes, the relationship between confidence intervals and acceptance criteria means that drugs with low intra-subject variability will have a greater chance of being declared bioequivalent than would otherwise be the case. The width of each 90 percent confidence interval is a function of the statistical power of the study, including the amount of intra-subject variability in the data. The more tightly grouped each test subject's measurements, the less variability there will be in the study data. This narrows the resulting confidence interval, increasing the likelihood that the interval will fall fully within the agency's 80 to 125 percent acceptance criteria.

This, in turn, means that test and reference products with low intra-subject variability may differ considerably and still be found bioequivalent. As illustrated in the graphic below, the narrower the confidence interval, the farther the point estimate (and the entire interval itself) can drift away from 100 percent (a test-to-reference ratio of 1.00) and still fit within FDA's acceptance criteria. As stated in the *Orange Book,* "[A] test product with low variability may pass the bioequivalence

criteria, when there are somewhat larger differences in the average response." *Orange Book* at x.

The following graphic shows the point estimates and confidence intervals for a hypothetical drug product ("A") with normal to high intra-subject variability. Figure A1 represents one manufacturer's version of the product; A2 represents another manufacturer's version. The graph also shows the point estimates and confidence intervals for a hypothetical drug product ("B") with low intra-subject variability. Significantly, the point estimates and confidence intervals for products B1 and B2 may drift farther than may products A1 and A2 from a ratio of 1.00, and still "pass" as bioequivalent.[20]



20 As discussed below, levothyroxine is a drug that demonstrates low intra-subject variability and, therefore, is more like hypothetical drug "B" in this graphic. *See infra* at 37.

In any case, an applicant's methodology – including study design and statistical analysis – must be shown to be capable of detecting clinically significant differences between a proposed drug and its reference drug. *See* General BA/BE Guidance at 6. If the test and reference drugs differ to a degree that is clinically significant, the methodology must be sensitive enough to prevent a claim of bioequivalence. Anything else and the agency runs the risk of declaring equivalent two products that, in fact, release different amounts of active drug into the body. *See generally* 21 CFR 320.24(a) (requiring that the method used to show bioequivalence must be the "most accurate, sensitive, and reproducible approach available . . . .").

Procedurally, in addressing the issue of BE methodologies for a class of products such as levothyroxine, FDA must consider two points. First, the agency is required by law to issue a public "guidance document" to communicate recommendations on the "testing of regulated products," except where the communication is directed to individual firms or persons. *Id.* at 10.115(b)(2), (b)(3). If the guidance involves "complex scientific issues," then the agency must publish a draft and seek public comment before finalizing the document. *Id.* at 10.115(g); *see* 21 USC 371(h). The agency may also, as part of this process, hold one or more public meetings to facilitate the development of the guidance. *See* 21 CFR 10.115(g). FDA is prohibited by law, however, from using any means – other than a guidance document – to communicate new or different regulatory expectations to a broad public audience. *See id.* at 10.115(e).

Second, the decision whether (or not) to refer the issues raised in this petition to an advisory committee must be well reasoned and consistent with agency standards and practice. While FDA enjoys discretion in its use of advisory committees, the agency must give interested persons a reasonable opportunity to have matters of scientific controversy reviewed "by an appropriate scientific advisory panel . . . or an advisory committee . . . ." *Id.* at 10.75(b)(2); *see* 21 USC 360bbb-1. This is particularly true of matters requiring "technical expertise." FDR Guidance at 7.

III. ARGUMENT

In 1997, FDA took decisive action to assert regulatory control over oral levothyroxine products. The intent was to ensure that thyroid patients would receive precise and consistent doses of levothyroxine each time they took the product, and each time they refilled their prescriptions. Today, with the approval and listing of seven brand-name levothyroxine products, the agency must again take action to ensure precise and consistent dosing. The key to ensuring that a generic levothyroxine product can be safely substituted for its brand-name counterpart is to

have in place a valid bioequivalence methodology that has been shown to meet the clinical needs of thyroid patients.

In January 2003, FDA informed Abbott that it had adopted, and was recommending to sponsors, a three pre-dose baseline correction method. This method, standing alone, has been shown to be insufficient to distinguish levothyroxine products that differ in bioavailability by 12.5 percent or more. In addition, the agency adopted this BE methodology without seeking expert advice from an advisory committee and without any form of public process, as required by law.

A solution to this critical public health issue is overdue. FDA must halt the review and approval of any additional generic levothyroxine products until it has initiated an appropriate public process to arrive at a valid methodology for establishing the bioequivalence and therapeutic equivalence of levothyroxine products.

A. FDA Must Establish a Clinically Sensitive and Scientifically Valid Bioequivalence Methodology for Levothyroxine Products

Levothyroxine is a narrow therapeutic range drug for which precise dosing is essential. As a first step in developing a valid BE methodology for levothyroxine products, the agency must consider the clinical evidence and determine the level of sensitivity and precision that is needed for purposes of declaring one sponsor's levothyroxine product substitutable for another sponsor's product. Only after the agency has determined the level of precision needed, from dose-to-dose and from refill-to-refill, can FDA design a valid BE methodology. That methodology must address both the need for baseline correction and the need for carefully chosen statistical criteria for assessing bioequivalence. As shown by Abbott's simulation study, and as confirmed by Study M02-417, the use of a pre-dose baseline correction method alone will not allow a sponsor to detect clinically relevant differences in the bioavailability of levothyroxine products.

1. FDA Must Consider the Relevant Clinical Issues

Since 1997, FDA has emphasized that small differences in the dose of levothyroxine products from refill-to-refill pose serious clinical risks. The agency must again focus on this issue – this time in the context of ensuring that the agency's recommended BE methodology is sufficiently sensitive to detect clinically relevant differences.

The agency's own conclusions on the need for precise dosing of levothyroxine products may be summarized as follows:

- In 1997, FDA reported that it had received 58 adverse drug experience reports associated with the potency of levothyroxine products. Nearly half of these occurred when patients refilled prescriptions for doses on which they had previously been stable. The agency emphasized, that for certain populations, even a small increase in dose may be hazardous. *See* 62 FR 43535; Tab 5 at 194.

- In its July 2000 medical review of the first approved levothyroxine NDA, FDA wrote that "it is critical to precisely titrate the dose of levothyroxine sodium to achieve and maintain the euthyroid state clinically and biochemically, thus avoiding the adverse consequences of under- and overtreatment . . . ." Tab 8 at 327.

- In its April 2001 denial of Knoll's citizen petition on the regulatory status of Synthroid®, the agency wrote: "The physician's reliance on the results of a TSH test to establish the optimal amount of replacement therapy is undercut when patients do not get the correct dose when filling and refilling their carefully calculated prescriptions." Tab 9 at 342.

- In a July 2002 memorandum to the Synthroid® NDA, FDA wrote: "Safe and effective titration requires availability of multiple dosage strengths that permit the full range of total daily dosages (e.g., 25-300 mcg) in increments of 12 or 12.5 mcg." Tab 20 at 673.

- In class labeling used by FDA for oral levothyroxine products, the agency recommends dosing adjustments of as little as 12.5 mcg for several groups of patients, including elderly patients with underlying cardiac disease and patients with severe hypothyroidism. *See* Tab 7 at 274.

Finally, the agency in 2001 prepared an analysis of the clinical issue illustrating why only a *nine percent* difference between levothyroxine products poses a serious health risk. The agency prepared this analysis in support of FDA's determination that Synthroid® is a "new drug" under section 201(p) of the FDCA. In a confidential appendix to the Knoll petition response, FDA argued – based on confidential and trade secret information – that patients who have been titrated to a specific strength of levothyroxine may suffer serious consequences if, with each refill, they receive a slightly different dose. According to the agency's analysis, a nine

percent difference (too low or too high) would be sufficient to cause adverse health consequences.[21]

FDA's conclusions regarding the therapeutic range of levothyroxine have been confirmed by clinical experts, several of whom appeared before the ACPS on March 13. These clinicians urged FDA to consider as part of any BE methodology the sensitivity of patients to fine differences in levothyroxine dosing. For example, Carlos Hamilton, M.D., presented the view of the American Association of Clinical Endocrinologists that "[d]osage changes of as little as 12.5 to 25 micrograms of oral thyroxine daily can, indeed, have significant effects on serum TSH and on the symptoms that our patients describe. These changes, whether they result from change in the dose or in the brand of thyroid hormone, can have important clinical effects on our patients . . . ." Tab 5 at 181.

For some patient populations, even smaller changes in the dose of levothyroxine can have damaging effects on the body. According to Mike Tuttle, M.D., of the Memorial Sloan Kettering Cancer Center, lifelong levothyroxine therapy is much like "chemotherapy" for thyroid cancer patients. *Id.* at 185.[22] Patients must be kept subclinically hyperthyroid, with enough levothyroxine in their bodies to suppress the thyroid gland, but not so much that they exhibit clinical symptoms of hyperthyroidism. *See id.* Consequently, these patients experience significant adverse effects from even slight under- or over-dosage. As Dr. Tuttle explained, "*very small changes in their dose, as little as missing one thyroid pill a week or taking one extra pill a week, can tip them over the edge into clinical thyrotoxicosis.*" *Id.* (emphasis added). A slight increase in dose can cause patients to develop rapid heartbeats, nervousness, and an inability to sleep; a slight underdose allows TSH to rise and creates a "risk for recurrence" of their cancer. *Id.*

Small and undetected changes in dose also can have a serious impact on children with congenital hypothyroidism. According to Rosalind Brown, M.D., Director of Clinical Trials Research at Children's Hospital in Boston, this condition is

---

21 This document was prepared by the agency on the basis of confidential commercial information, and was specifically designated by FDA as confidential. It has, therefore, not been included as an attachment. Portions of the confidential appendix quoted in Abbott's submissions to the agency during the formal dispute resolution proceeding have been redacted from the relevant attachments to this petition. *See* Tab 9 at 344.

22 Dr. Tuttle has in the past received grants from Abbott's predecessor, Knoll.

found in 1 out of every 3,000 infants. *See id.* at 183.[23] Babies born with congenital hypothyroidism face the risk of developing an IQ of less than 85, which represents significant cognitive impairment. Screening and levothyroxine treatment have made the disorder one of the most common treatable causes of mental retardation; however, it requires early and accurate hormone therapy. *See id.* Dr. Brown emphasized that "[r]elatively small differences in the dose of thyroxine replacement can have an enormous . . . and irreversible impact . . . in the outcome of these babies." *Id.*

Thyroid patients over age 60 also experience significant symptoms from the improper administration of levothyroxine. As explained at the ACPS meeting by Bryan Haugen, M.D. of the University of Colorado Health Sciences Center, older hyperthyroid patients have an increased risk of atrial fibrillation with only moderately suppressed TSH. *See id.* at 183-84.[24] These patients also experience reduced exercise capacity and cardiac function, decreased bone mineral density, a three- to four-fold increased risk of fracture, and increased all-cause mortality. *See id.* at 184. Even very minor adjustments in levothyroxine for these patients can have "dramatic effects." *Id.*

Dr. Haugen provided the committee with an example of these effects:

> A 62-year old woman presented with classic symptoms of hypothyroidism that you heard from Dr. Hamilton. She had fatigue, weight gain and constipation and her laboratory testing revealed a serum TSH that was elevated . . . at 28 and a serum T4 that was perfectly within the normal range, which many of us see in many different patients, and we call this mild thyroid failure or subclinical hypothyroidism.
>
> She was treated with 0.1 milligram of levothyroxine once a day. Eight weeks later, she returned. Symptoms had improved, still did have fatigue, and her serum TSH was still slightly elevated . . . at 7. Her serum T4 again was perfectly within the normal range and only slightly higher than her previous T4 of 8. The levothyroxine was increased by 25 micrograms, or 25 percent in this case, to 125 micrograms a day. Eight weeks later, her fatigue had somewhat improved, but now she had new insomnia, and as you can see, her TSH was now below the normal range at 0.08 milliunits per liter. . . .

---

[23] As far as Abbott is aware, Dr. Brown has no financial relationship with any companies whose products might be affected by this issue.

[24] As Dr. Haugen disclosed at the advisory committee meeting, he has in the past served as a consultant to Abbott.

> The levothyroxine was decreased to 112 micrograms per day, a decrease of only 10 percent. Seven weeks later, she returned with no complaints and her TSH now was in that target range we have talked about between 0.5 and 2.

Id. at 183-84.[25]

Finally, on April 4, 2003, the American Thyroid Association, in conjunction with the American Association of Clinical Endocrinologists, the Endocrine Society, and the Lawson Wilkins Pediatric Endocrine Society, sent a letter to Commissioner McClellan on the issue of levothyroxine bioequivalence. In it, the Association wrote that:

> Small differences between doses – *well within the range of formulation differences that might be undetected with the current bioequivalence method* – can have major clinical implications for thyroid patients, including symptoms, atrial fibrillation, osteoporosis, and uncontrolled hypercholesterolemia. As a result, we are extremely concerned that the current method used to determine levothyroxine bioequivalence and resulting wide tolerances for acceptability put patients at risk of adverse events.

Tab 21 at 677 (emphasis added).

These well-stated clinical concerns must inform the agency's decision on a recommended BE methodology for levothyroxine products. Clearly, a methodology that would allow for 12 or 13 mcg differences among "equivalent" products is inadequate. And, according to the agency's prior analysis, a nine percent difference – from refill-to-refill – can lead to serious adverse consequences. *See supra* at 24-25. Thus, the question the agency must address, based on the clinical evidence and on the agency's own prior analyses, is whether the BE methodology must be able to detect

---

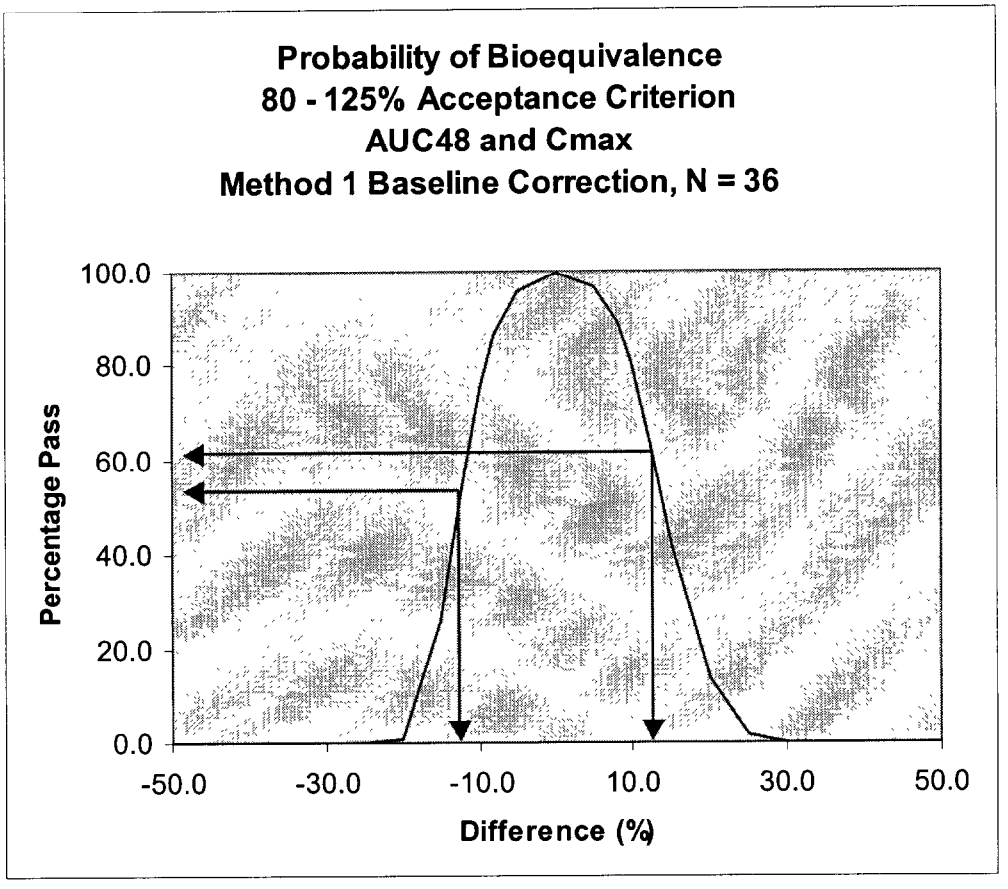
25 The only contrary view expressed at the March 13 ACPS meeting came from an agency official, who offered the following comment on one published study: "I think importantly, though, no significant differences were observed [in the study] in any clinical symptomology, weight, pulse rate or any clinical index over the range of thyroxine doses that were studied, 25 micrograms below or 75 micrograms above the optimal. . . . In other words, there's no connection between the TSH and the clinical observation." Tab 5 at 204. This is a remarkable statement; among other things, it is inconsistent with the consensus opinion of medical experts and the large body of literature supporting the need for 25 mcg and smaller dosing increments cited by the agency in support of its approval of Synthroid® and the other brand-name levothyroxine products. *See, e.g.*, Tab 8 at 328-34.

even smaller differences. Only after that question is answered can biopharmaceutics experts define a BE methodology that is adequate to the task.

2. FDA Must Take into Account the Impact of Endogenous T4

Once a clinically acceptable difference between substitutable levothyroxine products has been quantified, the agency must address the issue of baseline levels of endogenous levothyroxine. These baseline levels can introduce significant error into a BE study if they are not taken into account. Abbott's simulation study, conducted with data from the Synthroid® NDA, predicted that the methodology described in the BA Guidance could allow products that differ by 33 percent or more to be declared bioequivalent, and Abbott later confirmed this clinically in Study M02-417.

FDA, on the basis of Abbott's data, apparently agreed with Abbott that the BA Guidance was inadequate and, in response, adopted a three pre-dose baseline correction method. Abbott's Study M02-417 showed, however, that two otherwise identical doses of levothyroxine that differ by 12.5 percent (450 *versus* 400 mcg) cannot be distinguished under the BE methodology adopted by the agency. Further, following completion of Study M02-417, Abbott re-ran its stochastic simulation to predict the likelihood that two levothyroxine products that differ by 12.5 percent would be declared bioequivalent using the agency's recommended baseline correction method. The results of that analysis are shown below.



See Tab 13 at 612. Thus, in a 36-subject study, there would be a 52 (or 62) percent chance that a test product that delivers 12.5 percent less (or more) levothyroxine than a reference product will be declared by FDA to be bioequivalent to the reference product. *See id.* This level of risk is, simply put, unacceptable for a narrow therapeutic range drug such as levothyroxine.

At the March 13 ACPS meeting and during the formal dispute resolution proceeding, FDA made three arguments in defense of its choice to adopt a simple pre-dose baseline correction method. FDA argued: (1) That dosage form proportionality data from the approved levothyroxine NDAs support FDA's position; (2) that Abbott's data are biased by an unacceptably low "signal-to-noise" ratio; and (3) that Abbott used an insufficiently sensitive assay to measure T4 levels as part of its study. As discussed below, none of these arguments is persuasive.

a. FDA's demonstration of the adequacy of its pre-dose correction method is unpersuasive

To date, the only affirmative showing that FDA has made in favor of its adoption of the pre-dose correction method occurred at the March 13 ACPS meeting. In support of its decision to adopt this method, FDA showed the following slide, presenting dosage form proportionality data from four of the six brand-name products that had been approved as of the date of the meeting:

Total T_4 Adjusted for Baseline (Ratios of LSM – 90% Confidence Intervals)

[data from dosage-form equivalence studies]

Product	$AUC_{0\text{-}48\ hrs}$ A vs. B	$AUC_{0\text{-}48\ hrs}$ C vs. B	C_{max} A vs. B	C_{max} C vs. B
1	102.4% (94 7% - 110 8%)	100.2% (92 6% - 108 4%)	103 5% (97 3% - 110 0%)	97 7% (91 8% - 103.8%)
2	103 72% (95 98% - 112 09%)	91.45% (84.70% - 98 74%)	103 12% (96 87% - 109 76%)	95 05% (89 36% - 109 76%)
3	104% (97.09% - 110.35%)	98% (92 36% - 104.92%)	102% (94 94% - 108 57%)	100% (92 79% - 106 04%)
4	97% (90% – 105%)	114% (106% - 123%)	94% (87% - 101%)	104% (97% - 111%)

Treatment A = 12 x 50 mcg, Treatment B = 6 x 100 mcg; Treatment C = 2 x 300 mcg

24

Tab 5 at 217. This slide shows the results of FDA's retroactive application of the pre-dose correction method to the dosage form proportionality studies submitted under the BA Guidance.

That is, the dosage form proportionality studies conducted under the guidance were submitted without baseline correction. In all four cases, 600 mcg doses of twelve 50 mcg tablets and two 300 mcg tablets were each found to be bioequivalent (without correction) to a 600 mcg dose of six 100 mcg tablets. As the slide shows, each 600 mcg dosing regimen remained bioequivalent after the agency re-analyzed the data using pre-dose correction. In the circled cases, as emphasized by the agency, the 90 percent confidence intervals *with correction* came close to extending outside of the 80 to 125 percent acceptance criteria.

To FDA, this revealed the sensitivity of its new methodology: "We've got a 14 percent increase in . . . product 4, for AUC, and on the same scale, we also have

about a 9.5 percent decrease [in product 2]. The confidence limits, if this were slightly more variable, would have clearly failed." *Id.* at 199.

In fact, this slide provides no support for the agency's conclusion that its BE methodology is sufficiently sensitive. Critically, although the doses being compared purportedly contained the same amount of active ingredient, nothing is known about their true bioavailability. FDA's slide thus fails to support the agency's case because there is no analysis of whether or to what extent, for example, the two 300 mcg tablets and the six 100 mcg tablets, which passed as bioequivalent, might actually have differed from each other. Without controlling for any actual differences between the products, FDA has not reached the central issue raised by Abbott's data; that is, whether doses that *actually deliver* 12.5 percent more or less active ingredient would be found bioequivalent under FDA's methodology.[26]

b. FDA's "signal-to-noise" argument fails

As discussed above, the 450 and 400 mcg doses in Study M02-417 were shown to be bioequivalent using the same correction method that the agency has now adopted. As Abbott has argued, this finding demonstrates that the agency's BE methodology will result in the approval of some generic levothyroxine products that are not therapeutically interchangeable with their brand-name counterparts.

The agency, however, in the ODE II Letter, rejected Abbott's 450 to 400 mcg comparison on the ground that the ratio of exogenous "signal" (provided to subjects via study drug) to endogenous "noise" (the naturally-occurring levothyroxine present in the body of subjects) decreases as the size of the test dose decreases. Test doses that are "significantly below 600 mcg" will, according to CDER, result in too little signal and too much noise to yield accurate measurements. Tab 16 at 645. This reasoning, however, begs two key questions: (1) What is the scientific support for a 600 mcg dose; and (2) on what basis did the agency determine that 450 and 400 mcg doses are "significantly below" the level needed to yield accurate measurements? Without clearly articulated answers to both questions, the agency's position is unsupportable.

The only basis cited in the ODE II Letter for requiring a 600 mcg dose is the agency's levothyroxine BA Guidance. *See id.* There, the agency simply stated that "*several times the normal dose* should be given to raise the levels of the drug significantly above baseline to allow measurement." Tab 10 at 354 (emphasis added).

---

[26] In Study M02-417, by contrast, Abbott took several steps to control the actual differences in the bioavailability of the test doses. *See supra* at 11 n.12.

Later in the guidance, the agency recommended 600 mcg (*i.e.*, twice the *highest* available strength) as a suitable test dose. There was, however, no scientific showing by the agency that 600 mcg represented a critical threshold. No data were cited and no attempt was made to quantify or explain 600 mcg as the minimum necessary test dose. In fact, documents recently disclosed by the agency show that FDA accepted a BE study that used a 500 mcg test dose as one of the three studies in support of the only levothyroxine ANDA approved to date. *See* Tab 11 at 380.

Under the BA Guidance, doses even lower than 600 or 500 mcg would be appropriate. Again, the guidance recommends only "several times the normal dose . . . ." Tab 10 at 354. According to the labeling for this class of products, the average full dose of levothyroxine is approximately 1.7 mcg/kilogram ("kg"), or 100 to 125 mcg for a 70 kg adult. *See* Tab 7 at 273.[27] A test dose of 400 mcg would, then, be several times the normal dose.

If FDA's concern is about obtaining high enough levothyroxine levels to yield accurate measurements, the agency must acknowledge that the blood levels measured in Abbott's study do not differ markedly between a 600 mcg dose and a 400 mcg dose. The mean maximum T4 value after the 600 mcg dose in Study M02-417 was approximately 14 mcg/dl, and the mean maximum value after the 400 mcg dose was approximately 13 mcg/dl. *See* Tab 12 at 504. Importantly, all of these values are substantially higher than lower limit of quantification of the assay, and higher than the baseline of approximately 7.5 mcg/dl. *See id.* at 505; *see also infra* at 34-35.

FDA's reliance on a 600 mcg dose is even more suspect once one recognizes the higher levels of variability exhibited by the data derived from the 400 mcg and 450 mcg doses Abbott studied, as compared to the 600 mcg dose. As described above, the statistical analysis used to determine whether two products may be declared bioequivalent is performed on the logarithm of C_{max} and the logarithm of AUC. The two basic factors that determine bioequivalence are the relative bioavailability of the products and the variability of the data. The less variability in the data, the narrower the measures' confidence intervals and the more likely it is that two products will be found bioequivalent.

With Abbott's lower doses of levothyroxine, there was more variability in the data, resulting in wider confidence intervals than would have been the case at

---

27 *See also* IMS Health, *National Prescription Audit Plus* (Full Year 2002) (reporting that 100 to 125 mcg tablets represent approximately 40 percent of all prescriptions); Tab 11 at 371 ("The average full replacement dose of levothyroxine is 1.7 mcg/kg/day (e.g., 100-125 mcg/day for a 70 kg adult) for hypothyroidism in adults and in children in whom growth and puberty are complete.") (emphasis removed).

higher doses. Indeed, the variability of the data from the 600 mcg arm of Study M02-417 was smaller than the variability for the 450 and 400 mcg arms.[28] Thus, two levothyroxine products that differ by 12.5 percent are even *more* likely to be found bioequivalent in a study with 600 mcg doses than in a study with 450 and 400 mcg doses. The fact that Study M02-417 demonstrated bioequivalence at these levels is that much more compelling.

Furthermore, despite the greater degree of variability in the data at the 450 and 400 mcg levels, these two doses passed bioequivalence by a comfortable margin. The width of the confidence intervals in Study M02-417 for the 450 *versus* 400 mcg comparison under Correction Method 1 ranged from 0.14 to 0.25, relative to the width of the acceptance criteria, 0.45 (*i.e.*, from 0.80 to 1.25). *See* Tab 12 at 506. This suggests that doses that differ by even more than 12.5 percent, when tested at higher dose levels (providing less variability and narrower confidence intervals), would likely pass as bioequivalent as well.

Finally, CDER's reliance on a 600 mcg dose in the context of baseline corrected data is problematic, given that the agency originally recommended a 600 mcg dose in its BA Guidance for use *in lieu of* baseline correction. *See* Tab 10 at 354. The agency recommended the large, 600 mcg dose, to offset the impact of baseline T4 and ensure accurate measurement. *See id.* However, in criticizing Abbott's study, CDER fails to recognize that Abbott's data were corrected for baseline. Abbott showed that the 450 and 400 mcg doses could not be distinguished *after* the data were corrected for baseline – using the correction method now being recommended by FDA. In other words, CDER's primary basis for rejecting Abbott's argument has been that Abbott's test doses were too low; however, CDER has incorrectly judged Abbott's test doses against the dose FDA specifically selected for use *without* baseline correction.

---

28 The variances of the logarithms of C_{max} and AUC for the three dosing levels (*i.e.*, 600, 450, and 400 mcg) with the three pre-dose correction method were estimated, taking into account gender, dosing period, and unequal carryover effects. The estimates of the variances for the 600 mcg dose (0.0356 and 0.0336 for C_{max} and AUC, respectively) were smaller than for the 450 mcg dose (0.0563 and 0.0799) and the 400 mcg dose (0.0459 and 0.0574). These estimates were obtained by pooling the 12 corrected sums of squares from the 12 combinations of gender and sequence, and dividing this pooled sum of squares by (n-12).

In short, for CDER to assert that 600 mcg is better because it is higher, or that 400 mcg is unacceptable because it is too low, does not represent careful scientific analysis.[29]

c. FDA's assay sensitivity argument fails

At the March 13 ACPS meeting, the agency argued that the assay used to measure the levels of T4 in Abbott's clinical study was not sufficiently sensitive to detect differences between the doses that were tested. FDA, at the meeting, stated that "lower than [600 mcg], based on the data we had, we really did not think that anyone could really see the difference between formulations at a lower dose simply because of *lack of sensitivity of the assays* to even detect that in the blood." Tab 5 at 193 (emphasis added). The agency, however, offered no evidence or data in support of this position.

The assay that was used in Abbott's study, the DPC Coat-a-Count®, is an FDA-cleared radioimmunoassay for total T4. The performance characteristics of the assay were determined in a validation study, which demonstrated an analytical sensitivity, defined as the lower limit of quantification ("LLOQ"), of 1.00 mcg/dl. *See* Tab 12 at 502. Comparing the LLOQ of the assay to the values measured in Abbott's study, the lowest concentration of T4 for *any* subject in the study at *any* point (before or after dosing) was 4.11 mcg/dl. *See* Tab 12. Moreover, the mean peak levels in the study generally were in the 12 to 14 mcg range, again, well above the LLOQ. Thus, contrary to FDA's assertion regarding "lack of sensitivity of the assays to even detect that in the blood," all values observed in the study were well above the lower limit of reliable quantification.

---

29 Nor can the agency, as a matter of law, rely on the 600 mcg dose as a *de facto* minimum standard. *See Hoctor v. United States Dep't of Agriculture*, 82 F.3d 165, 170-71 (7th Cir. 1996) ("When agencies base rules on arbitrary choices they are legislating, and so these rules are legislative or substantive and require notice and comment rulemaking . . . ."). In *Hoctor*, the court invalidated a Department of Agriculture policy that perimeter fences around facilities housing dangerous animals should be at least eight-feet high. The court recognized the futility of trying to rebut a standardless numerical determination that the agency had adopted without explanation. As Judge Posner explained, "[t]here is no way to reason to an eight-foot perimeter-fence rule as opposed to a seven-and-a-half foot fence or a nine-foot fence or a ten-foot fence." *Id.* at 170. To the extent CDER has rejected Abbott's 450 and 400 mcg data, simply because those doses fell below the 600 mcg dosing level recommended in the BA Guidance, CDER is applying the BA Guidance as if it were a rule. Rules must be issued through a notice-and-comment process prescribed by law (5 USC 553); the application of a guidance, as if it were a rule, is a clear violation of the APA.

The LLOQ of Abbott's assay also compares favorably to the radioimmunoassay used by the sponsor of the only approved levothyroxine generic product. That assay, relied upon in the BE studies submitted in support of the sponsor's ANDA, apparently has an LLOQ of 16.025 nanograms per milliliter, or 1.60 mcg/dl. *See* Tab 22 at 707.[30] FDA accepted the results of these studies, despite the agency's stated concerns about assay sensitivity.

In addition to the fact that all measured values in Study M02-417 were well above the LLOQ, the assay provides adequate precision to detect meaningful differences in dose.[31] Furthermore, by using multiple measurements to create a serum concentration-time curve, the effect of any individual assay measurement error is reduced. Random variation in the error would yield approximately equal numbers of measurements above and below the "true" serum levels. When the total area under this curve is calculated, the net effect of any individual error is reduced, as positive and negative differences offset each other. Also, because AUC is reported as a mean value for multiple subjects, repetition of the assay among multiple subjects in an adequately powered study would further reduce the effect of any error. Thus, the calculation of AUC based on measurements at multiple points and across multiple subjects minimizes the theoretical effect of assay-related error.

In sum, FDA has not shown that the 450 *versus* 400 mcg comparison in Study M02-417 is invalid. FDA's arguments – regarding the dosage form proportionality data from the approved NDAs, the objection to test doses less than 600 mcg, and the perceived lack of sensitivity of Abbott's assay – have no merit. To be clear, Abbott is not arguing that the agency may or should recommend any particular test dose for BA or BE purposes. That decision is within the agency's discretion, subject to the procedural requirements discussed below. *See infra* at 41-44. Rather, Abbott is arguing that the 450 and 400 mcg data from Study M02-417, demonstrating the insensitivity of FDA's recommended BE methodology, are valid and sound.

---

30 The agency redacted information regarding assay sensitivity from its recently released review documents for the approved generic levothyroxine product. However, this information is available in public documents obtained from the New Jersey Drug Utilization Review Council. *See* Tab 22.

31 According to the assay's labeling, intra-assay precision is approximately 2.8 percent, and inter-assay precision is approximately 5.1 percent. *See* Tab 23 at 856. Abbott's own validation study, conducted during Study M02-417, confirmed these results. *See* Tab 12 at 502.

3. FDA Must Consider Adopting Tighter Statistical Parameters When Evaluating the Bioequivalence of Levothyroxine Products

During the March 13 ACPS meeting, an official from the Office of Generic Drugs characterized the sensitivity of FDA's standard BE methodology as follows:

> Generally, for most products with normal levels of variability, say CVs of 25 percent or as much as 30 percent, the mean data or the point estimates that we see in normal bioequivalence studies don't generally fall outside of 10 percent and most of them are around 3 percent either way because essentially the confidence interval has a width around that mean and it doesn't really take much movement away from center to cause the edge of that confidence interval to go over our limit and fail.

Tab 5 at 193.[32] This methodology is designed to ensure that the mean AUCs of generic drug products do not differ from their brand-name counterparts by more than 20 percent, based on a conclusion that, with very few exceptions, differences of up to 20 percent are clinically acceptable. *See* Tab 24 at 860; *see also* Guidance for Industry: *Statistical Approaches to Establishing Bioequivalence* (Jan. 2001) at 2 n.2.

FDA has consistently maintained, however, that it would modify these criteria for a specific product where such differences are not clinically acceptable. As the agency's Bioequivalence Task Force recognized in 1988:

> [S]ome drugs or drug classes may require tighter limits than the [then-existing 80 to 120 percent acceptance criteria]. These situations must be identified on the basis of clinical evidence demonstrating a need to tighten the generally applied standard. Such evidence could include, for example, a prospective clinical study demonstrating that the usual criteria for bioequivalence measurements are not stringent enough.

---

[32] These statistics are supported by two surveys conducted by FDA, cited in the *Orange Book*, that report average differences between the observed mean AUCs of brand-name and generic products of approximately ± 3.5 percent. *See Orange Book* at x. In the first survey, the average difference between the observed mean AUCs of brand-name products and generic products eligible for approval under early ANDAs was approximately ± 3.5 percent. About 80 percent of the differences were within ± 5 percent. *See* Tab 24 at 860. In the second, an examination of the 273 generic drug applications approved in 1997 revealed average observed mean differences of 3.47 percent for AUC (standard deviation 2.84) and 4.29 percent for C_{max} (standard deviation 3.72). *See id.* at 863.

Tab 25 at 893; *see also* General BA/BE Guidance at 20; Comments of Roger Williams, M.D., former Deputy CDER Director for Pharmaceutical Science, before the ACPS (Dec. 12, 1997) (stating that it would be a public health advantage to narrow the criteria for narrow therapeutic range drugs, such as warfarin, to prevent the marketing of products with 12 percent or greater differences in absorption).

Such is the case with levothyroxine sodium. The agency's current BE methodology may permit the approval of levothyroxine products that differ by 12.5 percent or more. By FDA's own analysis, differences of as little as 10 percent, nine percent, or less, can have clinical impacts on thyroid patients. *See supra* at 24-25.

Moreover, as discussed above, levothyroxine presents low intra-subject variability. In healthy volunteers, the absorption of levothyroxine varies little between test periods, in part because of the body's own tight regulation of hormone levels. The resulting measurements are therefore tightly grouped and create narrow confidence intervals; that is, there is a high degree of confidence that each subject's "true" value lies within this small range. From a biopharmaceutics perspective, these narrow intervals increase the likelihood that products varying by significant amounts may be declared bioequivalent; that is, the confidence intervals may drift farther from a test-to-reference ratio of 1.00 and still fit within FDA's traditional acceptance criteria. *See supra* at 20-21.

Finally, the application of the "standard" acceptance criteria (80 to 125 percent) in this instance would suggest that the agency is applying those limits as if they had been adopted as a rule. In fact, the agency has failed to follow the necessary procedures to require, expressly or implicitly, that all BE studies be analyzed against these limits. When an agency issues a pronouncement that has a "binding" effect on private parties or the agency, it is required by the APA to engage in notice-and-comment rulemaking. *See* 5 USC 553; *see also Croplife America v. EPA,* 329 F.3d 876, 881, 884 (D.C. Cir. 2003). This is true regardless of how the pronouncement is characterized – if the agency treats a statement as controlling, or if the agency leads private parties to believe their submissions will be declared invalid unless they comply, the statement is, "for all practical purposes, binding" and must be issued as a rule. *Appalachian Power Co. v. EPA*, 208 F.3d 1015, 1021 (D.C. Cir. 2000) (quotation removed). Given the weight of Abbott's evidence, the overwhelming clinical demand for precise dosing, and the fact that we are unaware of *any* previous instance in which FDA has narrowed its acceptance criteria, the agency's failure to narrow the criteria for levothyroxine would strongly suggest that it is applying the 80 to 125 percent criteria as if they have been set down in a rule.

The combination of levothyroxine's low intra-subject variability and narrow therapeutic index focuses unprecedented pressure on the statistical

methodology generally used to evaluate bioequivalence. By narrowing its 80 to 125 percent acceptance criteria in this case, FDA can ensure that the true bioavailability of approved levothyroxine generic products more closely matches that of the reference products. Failure to adopt this and other necessary revisions to the agency's recommended BE methodology means that generic levothyroxine products may vary by 9 percent, 12.5 percent, or more from their brand-name counterparts.

* * *

In sum, Abbott has shown in its simulation study and in Study M02-417 that a single-dose BE study in healthy subjects, using a pre-dose correction method and FDA's standard statistical criteria, cannot detect clinically significant differences between products. Although FDA accepted Abbott's simulation and Study M02-417 insofar as they demonstrated the need for baseline correction, the agency rejected Abbott's studies to the extent that they invalidated the baseline correction method that the agency chose to adopt. Because none of FDA's reasons for this rejection is well-founded, the agency must halt the use of this methodology. It should immediately initiate an appropriate public process and refer the issue of BE criteria for levothyroxine products to a joint clinical/biopharmaceutics advisory committee.

B. FDA Must Adopt a Bioequivalence Methodology through a Public Process

1. FDA Should Seek the Advice of a Joint Advisory Committee

The Food and Drug Administration Modernization Act of 1997 ("FDAMA") provides sponsors a statutory right to request advisory committee review of scientific disputes. *See* FDAMA 404 (codified at 21 USC 360bbb-1). In enacting this law, Congress recognized the need for greater dispute resolution procedures, and the utility of involving outside experts in the process. *See* H. Rep. 105-310 (Oct. 7, 1997) at 73 ("Neither the current law nor existing regulations provides an adequate basis for resolving scientific and medical disputes that arise in the course of FDA implementation of the law.").

In response, FDA amended its internal review regulations, adding the opportunity for a sponsor to request review of a "scientific controversy by an appropriate scientific advisory panel . . . or an advisory committee . . . ." 21 CFR 10.75(b)(2). As outlined in guidance, disputes involving "technical expertise that require some specialized education, training, or experience" generally should be referred to a committee. FDR Guidance at 7.

In light of these statutory and regulatory provisions, the agency's stated rationale for refusing to refer the issues raised by Abbott to a joint advisory meeting is inadequate. In the ODE II Letter, CDER denied Abbott's meeting request, concluding that EMDAC participation is unnecessary because the clinical issues are already "very clear to the Agency's own medical experts as evidenced by the BA Guidance" on levothyroxine products. Tab 16 at 645. CDER also stated that the March 13 ACPS meeting would be "sufficient" with respect to any outstanding technical issues. *Id.* at 646.

In fact, the BA Guidance on which CDER relied does not address the clinical issues implicated by the substitution of levothyroxine products. The background section of that document discusses generally the need for precise dosing, but does not address how closely matched interchangeable levothyroxine products must be. *See* Tab 10 at 354. As shown in Abbott's February 12 FDR Submission, there is a pressing need to consider this clinical issue specifically in the context of BE studies for products that will be considered fully interchangeable. *See* Tab 2 at 15-17.[33]

With regard to the technical biopharmaceutics issues, CDER's reliance on the March 13 meeting was, in retrospect, misplaced. As discussed above, on March 4, 2003, FDA published a revised advisory committee agenda that *withdrew* the specific topic of levothyroxine bioequivalence from the March 13 agenda. *See* 68 FR 10254. Several days later, FDA posted a public notice stating that the issue of Abbott's study and its impact on levothyroxine bioequivalence "*is not a topic for discussion at this ACPS meeting.*" Tab 17 at 656 (emphasis in original). Neither CDER's analysis nor the agency's slides presenting that analysis were provided in advance to Abbott or to the members of the ACPS. In addition, the agency did not present any questions to, or solicit any recommendations from, the ACPS on the issue.

In short, the issues raised in Abbott's formal dispute resolution proceeding, the March 13 ACPS meeting, and Study M02-417 and the related simulation studies involve clinical and technical issues that are particularly suited for review before clinical and biopharmaceutics experts.[34] Congress has directed the

---

[33] As recognized by an official from FDA's Office of Generic Drugs at the March 13 ACPS meeting, the purpose of therapeutic equivalence is to allow for drug substitution "without any adjustment in dose or other additional therapeutic monitoring." Tab 5 at 212.

[34] Although the precise makeup of such a joint advisory committee is beyond the scope of this petition, we note that the greatest concentration of technical expertise regarding these issues likely lies with the Clinical Pharmacology Subcommittee of the ACPS.

agency to use an "appropriate scientific advisory panel" to help resolve scientific disputes and the agency has committed that requests for advisory committee review to resolve scientific disputes "shall not be unreasonably denied." 21 USC 360bbb-1; 63 FR 63978, 63980 (Nov. 18, 1998). In this case, the need for advisory committee review is inescapable. Abbott has more than amply supported its request and, as the agency has recognized, this is a matter of significant public interest that should be decided in a public forum. *See* Tab 1.

In this light, Abbott proposes that the issues that must be presented to a joint advisory committee include:

- *Precision of dose required for interchangeable products.* Before designing any BE methodology for levothyroxine products, FDA must quantify the clinically acceptable difference that may be allowed between substitutable (*i.e.*, "AB" rated) levothyroxine products. The agency has already determined that a nine percent difference from refill-to-refill poses a serious health risk. *See supra* at 24-25.[35]

- *Baseline correction method.* FDA's methodology must include a method to address baseline levels of levothyroxine. *See supra* at 28-35. It must take into account error presented by several factors, such as the suppressive effect of exogenous levothyroxine on endogenous baseline levels, the diurnal variation of baseline levels, and the possibility of unequal carryover effects from prior test periods.

- *Modification of current statistical acceptance criteria.* As discussed above, FDA must consider narrowing the statistical acceptance criteria (currently 80 to 125 percent) to criteria that more closely meet the clinical demands of levothyroxine products and the known low intra-subject variability. *See supra* at 36-38.

---

35 In the context of determining clinically tolerable differences for BE purposes, the agency must also consider the clinical impact of the variation in drug content between the test and reference products. FDA generally requires that the drug content of the test and reference products used in bioequivalence tests not differ by more than 5 percent. *See* General BA/BE Guidance at 21. In the case of the only approved generic levothyroxine product, the agency accepted a *5.1 percent* difference. *See* Tab 11 at 386. In this instance, to ensure safe and effective substitution, FDA may need to *narrow* its 5 percent requirement or otherwise correct for this variation. *See, e.g.*, Health Canada Guidance for Industry: *Conduct and Analysis of Bioavailability and Bioequivalence Studies – Part A: Oral Dosage Formulations Used for Systemic Effects* (1992) at 19 (recommending correction of bioequivalence data for measured drug content), available at www.hc-sc.gc.ca/hpfb-dgpsa/tpd-dpt/bio-a_e.pdf.

- *Size of the test dose used.* FDA needs to determine how large a dose of levothyroxine sodium should be given to subjects, in relation to the method of baseline correction. The agency's current justification for reliance on a 600 mcg dose is arbitrary. *See supra* at 30-35.

- *Study population.* FDA must consider more carefully the possibility of conducting bioequivalence studies in patients, including athyreotic patients. Such studies could use therapeutic doses of the drug and would not require any baseline correction method.[36]

- *Washout period.* FDA has acknowledged the long (~1 week) elimination half-life of levothyroxine, and recommends a washout period of 35 days in the BA Guidance. Abbott's Study M02-417, however, demonstrated carryover effects of up to 53 days with supra-therapeutic doses. *See* Tab 12 at 508. FDA's new methodology may need to incorporate a longer washout period or another method of accounting for this effect.

- *Additional markers.* FDA may need to consider 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.

2. FDA Must Adopt Recommended Levothyroxine Bioequivalence Criteria in Compliance with GGPs

As discussed above, on May 15, 2003, CDER issued a letter requesting that Abbott submit a citizen petition on the issue of BE criteria for levothyroxine products. *See* Tab 1. In that letter, CDER acknowledged that this issue is of

---

[36] At the March 13 ACPS meeting, an agency official, when asked about the possibility of BE studies in athyreotic patients, stated that it was "unrealistic" due to recruiting difficulties and the lack of enough subjects. To the contrary, numerous studies have been conducted in this population, including several BE studies. *See, e.g.*, Tab 26 (attaching Shapiro, et al., *Minimal Cardiac Effects in Asymptomatic Athyreotic Patients Chronically Treated with Thyrotropin-Suppressive Doses of L-Thyroxine*, J. Clin. Endocrinol. Metab., vol. 82 (1997) at 2592 (involving 17 patients); Gottwald et al., *Bioequivalence of Two Commercially Available Levothyroxine-Na Preparations in Athyreotic Patients*, Meth. Find. Exp. Clin. Pharmacol., vol. 16 (1994) at 645-50 (24 patients); Trantow, et al., *A New Method for the Determination of the Bioavailability of Thyroid Hormone Preparations*, Meth. Find. Exp. Clin. Pharmacol., vol. 16 (1994) at 133 (24 patients); Mechelany, et al., *TRIAC has Parallel Effects at the Pituitary and Peripheral Tissue Levels in Thyroid Cancer Patients Treated with L-Thyroxine*, Clin. Endocrinol., vol. 35 (1991) at 123 (22 patients)).

"significant interest to other manufacturers of levothyroxine sodium, including generic drug applicants." *Id.* at 1. Therefore, CDER concluded that all interested persons should be given "the opportunity to participate in the decision-making process." *Id.* Beyond the manufacturers of these products, this issue is of critical interest to the broader thyroid community, as demonstrated by the April 4 letter from the leading endocrine medical societies to Commissioner McClellan (Tab 21), and by the numerous clinicians who spoke at the March 13 ACPS meeting (Tab 5 at 178-89).

Both Congress and FDA have established procedures for the development of agency policies regarding complex scientific matters of public concern. In the absence of notice-and-comment rulemaking, Congress articulated the procedural steps FDA must follow prior to issuing guidance. *See* 21 USC 371(h)(1)(C) (requiring public participation before the implementation of guidance concerning "complex scientific" or "highly controversial" issues). Congress further required the agency to issue regulations consistent with that practice. *See id.* at 371(h)(5); *see also* 21 CFR 10.115.

For guidance documents that set forth initial interpretations of regulatory requirements or deal with "complex scientific" or "highly controversial" issues, FDA's GGP regulations direct the agency to publish in the *Federal Register* an announcement that a draft guidance is available, accept and review comments, and publish a final draft before implementation. The agency also may hold public meetings or present the draft guidance to an advisory committee for review. *See* 21 CFR 10.115(g)(1); *see also id.* at 10.115(f)(7) (providing the same procedures for the revision of such guidance documents).

It is incumbent upon the agency to follow its GGP regulations to adopt a BE methodology for levothyroxine. Since 1997, the agency has made all significant announcements regarding levothyroxine through guidance documents because, according to FDA, it "recognized, in part due to the large number of manufacturers of this product, [the need] to come up with a consistent set of guidelines . . . ." Tab 5 at 194; *see supra* at 6. As of 1997, at least 37 sponsors were marketing levothyroxine products, and since then more than 10 have sought NDA and ANDA approval of levothyroxine products. *See* Tab 5 at 193.

FDA's recommendation of a levothyroxine BE methodology plainly qualifies as agency guidance. *See* 21 CFR 10.115(b)(1) (defining guidance documents as documents "prepared for FDA staff, applicants/sponsors, and the public that describe the agency's interpretation of or policy on a regulatory issue"); *see also id.* at 10.115(b)(2) (including as examples of guidance documents those relating to the testing of products and the evaluation of submissions). Nevertheless, the agency first made this policy known to Abbott in the Division Letter and, as far as Abbott understands, first publicly announced the guidance at the March 13 meeting of the

ACPS.[37] The public process required by the agency's GGP regulations has, to date, not been followed.

Moreover, the agency's current BA Guidance was never intended or designed to address the bioequivalence of levothyroxine products. *See id.* ("This guidance was intended to address issues of bioavailability . . . and was never intended to be used on its own for the purposes of bioequivalence."). Indeed, Abbott's simulation study and Study M02-417 show that its use in such a manner can result in the approval of generic products that differ from their reference products by 33 percent or more. FDA has nevertheless applied this document as a *de facto* BE guidance in at least one circumstance, compounding the impact of the agency's failure to follow GGPs.[38]

In adopting its baseline correction method without public process, FDA either issued new guidance on the subject of levothyroxine BE testing, or it revised the BA Guidance, without providing notice or accepting comment from interested persons. Either constitutes a violation of the agency's own GGP regulations. *See* 21 CFR 10.115(e) (prohibiting FDA from using means other than a guidance document "to informally communicate new or different regulatory expectations to a broad public audience for the first time"). Ultimately, the need for a BE guidance document is clear, and the need for a comprehensive clinical and statistical foundation for such a guidance is even clearer.

Given the significant public interest in this matter, FDA should not, and cannot, adopt a new BE methodology for this class of products by *ad hoc* means. By

---

37 In the Division Letter, the agency wrote that it "*will recommend to sponsors* seeking to obtain an AB rating of their product with respect to a reference listed levothyroxine sodium tablet product the following: *It will be necessary* to conduct a . . . study . . . using a . . . baseline subtraction method . . . ." Tab 4 at 151 (emphasis added).

38 In a memorandum to the ANDA file of the only approved generic levothyroxine product, FDA wrote the following:

> The bioequivalence criteria are calculated using data that is not baseline corrected based upon *current agency policy* regarding this specific drug product. This policy is outlined in the Guidance to Industry [sic] *Guidance for Industry, Levothyroxine Sodium Tablets – In Vivo Pharmacokinetic and Bioavailability Studies and In Vitro Dissolution Testing* issued December 2000.

Tab 11 at 477 (first emphasis added); *see* General BA/BE Guidance at 4 ("[W]e recommend that similar approaches to measuring BA in an NDA generally be followed in demonstrating BE for an NDA or an ANDA.").

issuing its recommendation without following GGPs, FDA deprived itself of the benefit of public input; in doing so, the agency appears to have adopted a scientifically flawed approach. FDA must now correct these procedural violations by initiating a public process – consistent with GGPs and including a joint advisory committee meeting – to develop a BE methodology that is sensitive enough to detect clinically significant differences between products. For these reasons, the agency also must halt the review of any applications that rely on FDA's current – unlawful – guidance.

IV. CONCLUSION

The agency is required by law to refuse to approve any applications for oral levothyroxine products that turn on a showing of bioequivalence where the methodology has not been shown to be accurate, sensitive, and reproducible. Among other things, any application that relies on FDA's current BE methodology must be refused approval, because such a method lacks the sensitivity needed to detect clinically relevant differences between products.

Further, Abbott respectfully requests that FDA initiate a public process to establish a BE methodology for oral levothyroxine products. This process should include a joint ACPS and EMDAC advisory committee meeting devoted specifically to developing a valid BE methodology for levothyroxine drug products. Following such a meeting, a proposed methodology must be presented for public comment under, at minimum, FDA's good guidance practice regulations. It is critical to patient health and safety that a methodology not be adopted or applied until the public has an opportunity to participate in the decision-making process. There is simply no other way to ensure the adoption of a clinically sensitive and scientifically valid BE methodology for levothyroxine products.

ENVIRONMENTAL IMPACT

The actions requested in this petition are subject to categorical exclusions under 21 CFR 25.30 and 21 CFR 25.31.

ECONOMIC IMPACT

Information on the economic impact of this proposal will be submitted upon request of the Commissioner.

CERTIFICATION

The undersigned certifies that, to the best knowledge and belief of the undersigned, this petition includes all information and views on which the petition relies, and that it includes representative data and information known to the petitioner which are unfavorable to the petition.

John M. Leonard, M.D.
Vice President
Global Pharmaceutical Development
Global Pharmaceutical Research and Development
Abbott Laboratories
(847) 938-4545

Douglas L. Sporn
Divisional Vice President
Regulatory Affairs
Global Pharmaceutical Research and Development
Abbott Laboratories
(847) 937-7986

Neal B. Parker
Senior Counsel
Legal Division
Abbott Laboratories
(847) 937-5201

Of Counsel:
David M. Fox
Brian R. McCormick
Hogan & Hartson L.L.P.
(202) 637-5600

cc: FDA Docket No. 03P-0126 (Citizen Petition of Jones Pharma, Inc.)