



Authors' Reply to Nicolas: "Why Were More than 200 Subjects Required to Demonstrate the Bioequivalence of a New Formulation of Levothyroxine with an Old One?"

Didier Concordet¹ · Peggy Gandia¹ · Jean-Louis Montastruc² · Alain Bousquet-Mélou¹ · Peter Lees³ · Aude A. Ferran¹ · Pierre-Louis Toutain^{1,3} 

Published online: 5 December 2019
© The Author(s) 2019

We welcome the opportunity to respond to Dr. Nicolas' comments [1] on our article in *Clinical Pharmacokinetics* [2]. On one point we can agree, however. The public domain headlines, on what *he* describes as the "Levothyrox[®] scandal", were indeed sensational, but also proportionate. Beyond this, there is little agreement between us and we therefore respond to his "challenge" (his choice of word) to our "argumentation". First, however, we make the general observation that he does not serve his own case well, when he takes our opinions out of context, to the point, in some instances, of distortion. Readers of *Clinical Pharmacokinetics* can be the judge of our two articles, Dr. Nicolas' challenge to them and this response to his challenge.

He claims wrong reasoning (on our part) in applying to individuals the tighter acceptance range (0.90–1.11) validated only for average bioequivalence (ABE) and he alleges that we are promoting individual bioequivalence (IBE). We continue to hold to the view that it is wrong to use the results of an ABE trial to support substitution of a new for an old formulation, which, moreover, *imposed* the former on more than 2 million patients. It is irrefutable that a planned ABE

trial incorporating more than 200 subjects implicitly reflects a large within-subject variability (WSV). This is incompatible with the necessity of prudently using a drug having a narrow therapeutic index, as is the case for levothyroxine (T4). Moreover, this is not consistent with the current marketing, in France, of 11 levels of strength of the new formulation of Levothyrox[®], intended to allow a very fine tuning of the prescribed dose and its possible adjustment.

Nicolas is correct in that we questioned, on scientific grounds, an ABE trial involving more than 200 subjects [2] but we did not and do not claim to make any moral or ethical judgment on this trial. At no time have we said that the Company wished to "*intentionally mask*" a large WSV. On the contrary, we stated clearly that this trial was both technically and professionally conducted [3]. To be clear, however, what we do challenge is the very principle of using an ABE trial to support a compulsory switch of an old formulation (OF) with a new formulation (NF) of Levothyrox[®], then *imposing* this change on millions of patients [3]. It is neither unfair nor incorrect to point out, and we now repeat that, whatever the WSV, ABE can always be demonstrated, simply by increasing the number of enrolled subjects, when the average $\mu T/\mu R$ ratio is equal or close to 1. This is not a matter of debate, it is self-evident, and we refute Dr. Nicolas' comment that it is "*unfair, casting a detrimental doubt in patient and public opinion*" (sic).

We continue to advocate and prioritize the *conceptual framework* of IBE because it places the patient firmly at the center of what should be the pre-occupation of all of us. Our argument goes well beyond simply regarding ABE as an in vivo quality control of two formulations, as pointed out by others [4]. We have not and we do not suggest, let alone require, the realization of an IBE trial, as historically promoted by the US Food and Drug Administration (FDA). This we made clear in our first publication when stating, "*IBE has been both extensively discussed and challenged*

This reply refers to the original article available at <https://doi.org/10.1007/s40262-019-00812-x>.

This reply refers to the comment available at <https://doi.org/10.1007/s40262-019-00849-y>.

✉ Pierre-Louis Toutain
pltoutain@wanadoo.fr

¹ INTHERES, Université de Toulouse, INRA, ENVT, 23 Chemin des Capelles, 31076 Toulouse, France

² Service de Pharmacologie Médicale et Clinique, Faculté de Médecine, Centre Hospitalier Universitaire de Toulouse, Université de Toulouse, Toulouse, France

³ The Royal Veterinary College, University of London, London, UK

and then, finally, not adopted by regulatory authorities. It is beyond the scope of this paper to discuss in detail advantages and limitations of IBE". On the contrary, as indicated in our second article, we suggest that the current FDA approach of assessing bioequivalence (BE) of levothyroxine formulations [5–7], comparing not only the average T4 area under the plasma concentration–time curve (AUC) and maximal plasma concentration but also the WSV of the two formulations using a replicate design, is the best approach. This is because the FDA guideline on levothyroxine requires documenting what the patient needs to know, namely a guarantee on the reproducibility of treatment. Essentially, this FDA approach is still an average but extended ABE and not an IBE.

Our previous article stated, and we re-iterate as a matter of incontrovertible fact, that the pivotal ABE trial conducted by the company reported a residual of 23.7%. Therefore, we do not endorse the conclusion of Nicolas that the Levothyrox[®] NF has “a moderate WSV but possibly lower than that of the old formulation Levothyrox[®] OF”. This is simply because, when the company planned the required number of subjects to conduct its pivotal ABE trial, they considered the already known intra-individual variability of T4 for the OF—estimated to be 11.54 or 15.3% for the variable AUC [8]. These figures are much lower than 23.7%, which in a 2 × 2 cross-over design reflects the WSV of both the OF and NF and which are assumed to be equal. It is a problem for us, but apparently not for Nicolas, to believe that a lower than 15% variability for the NF would lead to a 23.7% variability, when merged with that of the OF estimated at between 12% and 15%, because the corresponding computable probability of such a possibility is close to 0.

In the dose-proportionality study, the company reported a WSV of 17.1% for AUC [9] but this trial cannot be used to estimate unequivocally the WSV of the NF. This is because, under the experimental conditions of this trial, administration of up to 12 tablets of 50 µg strength (group A) was compared with administration to the same subjects of three tablets containing 200 µg (group C). This leads ineluctably to pharmaceutical bio-inequivalence between the two groups, by markedly altering the ratio of active ingredient to total vehicle (i.e., the same total dose for both groups but a four-fold higher amount of the vehicle for group A).

We continue to assert that a relationship exists between the value of WSV and the proportion of subjects possibly outside the a priori BE range, whatever its values. Moreover, we are not mistaken, as Nicolas alleges, when considering that this point has been properly documented by others, demonstrating that, in as many as 60% of individuals in a cross-over trial conducted in 24 subjects, the difference between parameters for the two formulations could, on average, be outside the range of BE (0.7–1.3 for this simulation) and yet still satisfy the regulatory criterion of ABE [10]. That this is

not merely an academic exercise was shown recently by the results of 14 four-sequence cross-over studies in volunteers receiving a range of drugs [11]. In this large set of trials, involving 700 subjects, it was reported that the percentage of individual AUCs outside the 0.8–1.25 a priori BE range was 16% on average and ranged from 2% for cephalexin to 35% for atenolol and clarithromycin. For maximal plasma concentration, the average was 32%, with values ranging from 8% for metronidazole up to 57% for diclofenac [11].

We do agree with Nicolas that it is unclear how the mean of 9.3% reported by the FDA was computed for the residual variability of levothyroxine from Abbreviated New Drug Application (ANDAs) reviewed between 1996 and 2008 [12]. However, as this figure was released in a tutorial workshop, we have assumed that it was retrospectively computed by the FDA with a baseline correction. More recently, the FDA approved a new formulation of levothyroxine sodium (Tirosint[®]-SOL, NDA 20697) [13] for which the sponsor conducted two BE trials using a three-way cross-over design, in which the WSV of this new oral formulation was calculated. For AUC after a baseline correction, the WSV was estimated to be 11.5% (study 140143) and 8.99% (study 140161). These clearly low values indicate that a baseline correction does not, either unduly or fatally, inflate the WSV of a well-developed levothyroxine formulation. What *does* inflate a WSV is a non-reproducible formulation and, once again, it should be emphasized that the baseline correction is just made to reveal the relevant variability to be considered for a BE trial.

Finally, we respond to Nicolas on the question of the subject-by-formulation interaction (SBFI). We did not state that a possible bio-inequivalence of the Levothyrox[®] NF with Levothyrox[®] OF must be necessarily ascribed to such a SBFI. What we did was to discuss the two possible explanations for a possible individual bio-inequivalence and their differing consequences, namely a SBFI requiring only a dosage adjustment for a sub-group of subjects and a large WSV, which, more significantly, would require a fundamental reconsideration of the formulation itself. Nicolas reported no SBFI in nine replicate designs [14]. This is not surprising when the consideration is successful BE trials, but absence of proof is not proof of absence. When the US FDA attempted to promote IBE, they reviewed their own data sets and concluded that many of the sets examined by their Center for Drug Evaluation and Research had estimated large SBFI values [15]. We comment that SBFI has been the single most controversial component of IBE. To conclude, we respond to the perhaps condescending but certainly biased Nicolas' comment on the connection between ABE and IBE, as reported by Munk [16]. In this publication, it is clearly indicated that the relationship between ABE and IBE “depends heavily on the underlying sample size” and “when sample size increases to infinity, this relationship vanishes”.

In discussing this issue, the author used $n = 12$ as a sample size, not 204, the number of subjects enrolled in the Levothyrox[®] ABE study. Most importantly, Munk concluded “Therefore, we come to the conclusion that at this ‘state of the art’ the standard test TOST (i.e., ABE) combined with a test for the equivalence of the within-subject variances represents a reasonable practical solution”. This is precisely what is now recommended by the US FDA guideline and with this we concur.

Compliance with Ethical Standards

Funding No sources of funding were used in the preparation of this reply.

Conflict of interest Didier Concordet, Peggy Gandia, Jean-Louis Montastruc, Alain Bousquet-Mélou, Peter Lees, Aude A. Ferran, and Pierre-Louis Toutain have no conflicts of interest that are directly relevant to the contents of this reply.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Nicolas P. Comment on: “Why were more than 200 subjects required to demonstrate the bioequivalence of a new formulation of levothyroxine with an old one?”. *Clin Pharmacokinet*. 2019. <https://doi.org/10.1007/s40262-019-00849-y>.
- Concordet D, Gandia P, Montastruc J-L, Bousquet-Mélou A, Lees P, Ferran AA, et al. Why were more than 200 subjects required to demonstrate the bioequivalence of a new formulation of levothyroxine with an old one? *Clin Pharmacokinet*. 2019. <https://doi.org/10.1007/s40262-019-00812-x>.
- Concordet D, Gandia P, Montastruc J-L, Bousquet-Mélou A, Lees P, Ferran A, et al. Levothyrox[®] new and old formulations: are they switchable for millions of patients? *Clin Pharmacokinet*. 2019;58:827–33. <https://doi.org/10.1007/s40262-019-00747-3>.
- Morais JAG, Lobato Mdo R. The new European Medicines Agency guideline on the investigation of bioequivalence. *Basic Clin Pharmacol Toxicol*. 2010;106(3):221–5.
- Anonymous. Draft guidance on levothyroxine sodium. Food and Drug Administration; 2018: p. 2. <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM428208.pdf>. Accessed 4 Oct 2019.
- Yu L, Jiang W, Zhang X, Lionberger R, Makhlof F, Schuirmann D, et al. Novel bioequivalence approach for narrow therapeutic index drugs. *Clin Pharmacol Ther*. 2015;97(3):286–91.
- Jayachandran P, Okochi H, Frassetto LA, Park W, Fang L, Zhao L, et al. Evaluating within-subject variability for narrow therapeutic index drugs. *Clin Pharmacol Ther*. 2019;105(2):411–6.
- Anonymous. An open-label, single-dose, randomized, two-period, two-sequence crossover, single-center trial to assess bioequivalence of 600 µg levothyroxine new formulation versus old formulation administered orally as 3 white tablets of 200 µg in healthy volunteers. 2014. https://ansm.sante.fr/var/ansm_site/storage/original/application/3606e51c6c64f05e25565a3c5bf840b7.pdf. Accessed 26 Nov 2019.
- Gottwald-Hostalek U, Uhl W, Wolna P, Kahaly GJ. New levothyroxine formulation meeting 95–105% specification over the whole shelf-life: results from two pharmacokinetic trials. *Curr Med Res Opin*. 2017;33(2):169–74.
- Endrenyi L, Schulz M. Individual variation and the acceptance of average bioequivalence. *Drug Inf J*. 1993;27(1):195–201.
- Hammami MM, De Padua SJS, Hussein R, Al Gaai E, Khodr NA, Al-Swayeh R, et al. Generic-reference and generic-generic bioequivalence of forty-two, randomly-selected, on-market generic products of fourteen immediate-release oral drugs. *BMC Pharmacol Toxicol*. 2017;18(1):78.
- Yu L. Quality and bioequivalence standards for narrow therapeutic index drugs. GPhA 2011 Fall Technical Workshop. FDA; 2011. <https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/UCM292676.pdf>. Accessed 26 Nov 2019.
- Anonymous. Tirosint-SOL (levothyroxine sodium) oral solution. 2016. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/206977Orig1s000TOC.cfm. Accessed 26 Nov 2019.
- Yu Y, Teerenstra S, Neef C, Burger D, Maliepaard M. A comparison of the intrasubject variation in drug exposure between generic and brand-name drugs: a retrospective analysis of replicate design trials. *Br J Clin Pharmacol*. 2016;81(4):667–78.
- Hauck WW, Hyslop T, Chen ML, Patnaik R, Williams RL. Subject-by-formulation interaction in bioequivalence: conceptual and statistical issues. FDA Population/Individual Bioequivalence Working Group. Food and Drug Administration. *Pharm Res*. 2000;17(4):375–80.
- Munk A. Connections between average and individual bioequivalence. *Stat Med*. 2000;19(20):2843–54.