



## Authors' Reply to Krebs-Brown et al. Comment on: "Why Were More Than 200 Subjects Required to Demonstrate the Bioequivalence of a New Formulation of Levothyroxine with an Old One?"

D. Concordet<sup>1</sup> · P. Gandia<sup>1</sup> · J. L. Montastruc<sup>2</sup> · A. Bousquet-Mélou<sup>1</sup> · P. Lees<sup>3</sup> · A. A. Ferran<sup>1</sup> · P. L. Toutain<sup>1,3</sup> 

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

As a preliminary response to Krebs-Brown et al. [1], we remind readers that *Clinical Pharmacokinetics* does not publish 'opinions'. Rather, in a section of the journal entitled "Current Opinion", it publishes regular articles and these are subjected, like all articles accepted by the journal, to a peer-review evaluation. We also remind readers that the bioequivalence (BE) trial that we have commented on has itself been published in a journal entitled *Current Medical Research and Opinion* (Gottwald-Hostalek, Uhl, Wolna, & Kahaly, 2017). As a first comment, we ask readers to note that we have *not* downgraded this BE trial to the status of a Merck-Serono opinion, despite the title of this journal.

However, the Letter to the Editor from Krebs-Brown et al. does offer an opinion, under the guise of seeking "further clarification". The five co-authors of the Krebs-Brown et al. "comment" are employees of the company that developed and successively marketed the old and new formulations of levothyroxine, on which we have commented in our two previous articles in *Clinical Pharmacokinetics*. We take it that the opinions expressed by Krebs-Brown et al. are those also of their employer, or at least sanctioned by the company. We are pleased now to answer the Merck Serono employees'

opinions. In seeking to denigrate our balanced conclusions with their opinions, they fail to answer both the arguments we have put forward and conclusions reached. Inappropriate and confrontational as they may be, we choose to pass over the comments of Krebs-Brown et al. "Without wanting to appear petty, we should point out that the paper contains several inaccuracies and misrepresentations ... we would like to distance ourselves from .... This implies that sponsors and planners of ABE trials inherently display a callous attitude to participants in such a trial, an accusation we must strongly protest". We choose rather to focus on the underlying science on which our conclusions are based.

We re-affirm our previous comments on the question of switchability. This was the key issue to stimulate debate, when replacement of the old formulation (OF) by a new formulation (NF) of Levothyrox<sup>®</sup> was imposed on millions of patients. Now, we are pleased to note that Merck Serono acknowledges that "it is unclear how to interpret the percentage of observed individual exposure ratio (IER) outside of this range" (i.e. the a priori BE range). Krebs-Brown et al. do not deny, in their response, that there was in their trial a large number of subjects outside the a priori BE range. Of significance is their acknowledgement that they are unable to interpret their own average BE (ABE) results on the question of switchability. Whilst there is indeed no regulatory recommendation on this point, this is not sufficient reason to simply ignore it.

This question of the number of individual exposure ratios outside the a priori BE range, as a possible weakness of an ABE study, has been addressed both very early and subsequently in the long history of BE. Using simulations, it was shown, in a cross-over trial conducted in 24 subjects, that as many as 60% of the individuals can, on average, be outside the range of BE (0.7–1.3 for this simulation) and yet still satisfy the regulatory criteria for ABE [9]. That this is not merely a speculative academic exercise has been demonstrated recently by the results of 14 four-sequence cross-over

---

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/s40262-019-00847-0>.

---

✉ P. L. Toutain  
pltoutain@wanadoo.fr

<sup>1</sup> INTHERES, Université de Toulouse, INRA, ENVT, Toulouse, France

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

<sup>3</sup> The Royal Veterinary College, University of London, London, UK

studies for a range of drugs [11]. In this large set of trials, involving 700 subjects, it was reported that the percentage of individual area under the curve (AUC) values 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 maximum plasma concentration, the average was 32%, with values ranging from 8% for metronidazole up to 57% for diclofenac [11]. This structurally and experimentally proven weakness of the ABE design becomes, in our opinion, a medical issue for drugs such as levothyroxine, classified as narrow therapeutic index drugs.

It does not suffice for Krebs-Brown et al. to offer to us a lesson in ethics regarding patients participating in an ABE trial. They must address our concern in seeking to evaluate the BE of levothyroxine in the *conceptual framework* of individual BE, even if conducting an individual BE trial (as historically proposed by the US Food and Drug Administration) has not been adopted by regulatory authorities. The attractiveness of the individual BE concept, notably in the present context, is that it places the patient and his/her expectations firmly at the heart of the trial, by considering his/her individual therapeutic window [6]. In contrast, an ABE has been meaningfully discussed by others [12] commenting on the current 2010 European Medicines Agency guideline as follows: “In fact, those parameters (i.e. AUC and maximum plasma concentration) seem to be more sensitive to differences in the formulation or the manufacturing process than clinical end-points and a more ‘quality-like’ approach has been adopted (in this guideline)”. We are amongst those who refuse to consider that, for a narrow therapeutic index drug such as levothyroxine (and we insist on so classifying levothyroxine), patients are simply members of a statistical distribution, for which it is sufficient to guarantee the geometric mean (or median) to fulfil their legitimate expectations to be treated with a reproducible formulation.

We do not doubt that Merck Serono knows exactly how to compute an a priori sample size to conduct an ABE. For the ABE here in question, the initial number of planned subjects was 216 [10]; a very large number. We re-iterate the basis of why we are challenging such a high number. First, 40 subjects is an acceptable number to cope with the risk of ‘drop-outs’ associated with a long wash-out interval. There remain 176 subjects that have been computed, using standard preliminary information, namely the two statistical risks alpha and beta. Alpha is fixed by regulation to 5% and we have assumed that the company wished a high power for its trial, i.e. 90%. We also considered a possible deviation of 5% in exposure between the two formulations. The remaining component to be considered for this computation, *not reported by Merck Serono*, is the within-subject variability (WSV). The WSV can be retrospectively estimated to approximately 17%, which is the value reported for their dose-proportionality trial [10]. What is disappointing here is

the anticipation of a rather large WSV for this NF, a formulation intended to be an improvement on the OF. Actually, the WSV common for the OF and NF was estimated in the pivotal trial under consideration to be 23.7% for AUC [10].

In contrast, it is reported in a document made public by the French authorities that the WSV of the OF was previously estimated to be 11.54 or 15.3% for AUC (Anonymous [4]). In other words, the NF is very unlikely to be an improvement on the OF in terms of reproducibility, a critical property for a NTI drug. Indeed, using standard computation on known variances of the OF and that of the pivotal ABE, Merck Serono employees are in position to estimate an order of magnitude of the WSV of their NF. To release into the public domain this estimate would allow determination, or not, of whether their NF has a reproducibility in line with the logic of their company in marketing 11 levels of strength of scored tablets to provide a prescriber with an exceptionally good level of dose adjustment for their patients. Our conclusion is that the major information missing from this dossier is an experimental estimate of the WSV of the NF that can be achieved with a replicate design, as currently proposed by the US Food and Drug Administration for levothyroxine (Anonymous [5] and which we are supporting [8]).

Once again, we repeat here the claim made in our first article [7], namely that “there is neither conspiracy nor malice on either side of this debate, but rather a difference of judgement on data derived and conclusions drawn”. It is essential that scientific argument should fuel this crucial debate on future developments, taking cognisance of the fact that EU scientific guidelines “do not have legal force and the definitive legal requirements are those outlined in the relevant Community legislative framework (Directives, Regulations, Decisions ...)”. The European Union guidelines are a non-binding consensus document “that applicants shall take into account” (Anonymous [2]). In other words, the European Union guideline on BE can be adapted in the best interest of patients. Furthermore, prior to submission of documentation, a company *should seek scientific advice, to discuss any proposed deviations during medicine development*. As acknowledged by the European Medicines Agency, the development of product-specific guidance, based on the outlined general principles (Anonymous [3]), is now timely. In this proposed review, levothyroxine should be a priority for the European Medicines Agency, exactly as was the case for the US Food and Drug Administration, when they proposed valuable guidance specifically for levothyroxine.

To conclude, we remind readers of *Clinical Pharmacokinetics* that the issue of overriding importance in this debate must be that human element, which comprises some millions of patients [13], who have or may have been involved in the switch from old to new formulations, a switch that was *imposed* on them.

## Compliance with Ethical Standards

**Funding** No sources of funding were received for the preparation of this reply.

**Conflict of interest** D Concordet, P. Gandia, J.L. Montastruc, A. Bousquet-Mélou, P. Lees, A.A. Ferran and Pierre-Louis Toutain have no conflicts of interest that are directly relevant to the content 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

- Krebs-Brown A, Munafo A, Gaikwad S, Urgatz B, Castello-Bridoux C. 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/s40262-019-00847-0>.
- Anonymous. Status of EMEA scientific guidelines and European pharmacopoeia monographs and chapters in the regulatory framework applicable to medicinal products. 2008. [https://www.ema.europa.eu/en/documents/scientific-guideline/status-emea-scientific-guidelines-european-pharmacopoeia-monographs-chapters-regulatory-framework\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/status-emea-scientific-guidelines-european-pharmacopoeia-monographs-chapters-regulatory-framework_en.pdf). Accessed 21 Nov 2019.
- Anonymous. Concept paper on the development of product-specific guidance on demonstration of bioequivalence. 2013. [https://www.ema.europa.eu/en/documents/scientific-guideline/concept-paper-development-product-specific-guidance-demonstration-bioequivalence\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/concept-paper-development-product-specific-guidance-demonstration-bioequivalence_en.pdf). Accessed 21 Nov 2019.
- 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](https://ansm.sante.fr/var/ansm_site/storage/original/application/3606e51c6c64f05e25565a3c5bf840b7.pdf). Accessed 21 Nov 2019.
- Anonymous. Draft guidance on levothyroxine sodium (p. 2). 2018. <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM428208.pdf>. Accessed 4 Oct 2019.
- Chen ML, Patnaik R, Hauck WW, Schuirmann DJ, Hyslop T, Williams R. An individual bioequivalence criterion: regulatory considerations. *Stat Med*. 2000;19(20):2821–42.
- Concordet D, Gandia P, Montastruc J-L, Bousquet-Mélou A, Lees P, Ferran AA, et al. Authors' reply to Coste et al.: "Levothyrox<sup>®</sup> new and old formulations: are they switchable for millions of patients?". *Clin Pharmacokinet*. 2019;58(7):967–8. <https://doi.org/10.1007/s40262-019-00780-2>.
- 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> (Epub ahead of print).
- Endrenyi L, Schulz M. Individual variation and the acceptance of average bioequivalence. *Drug Inf J*. 1993;27(1):195–201. <https://doi.org/10.1177/009286159302700135>.
- 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. <https://doi.org/10.1080/03007995.2016.1246434>.
- 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. <https://doi.org/10.1186/s40360-017-0182-1>.
- Morais JA, Lobato Mdo R. The new European Medicines Agency guideline on the investigation of bioequivalence. *Basic Clin Pharmacol Toxicol*. 2010;106(3):221–5. <https://doi.org/10.1111/j.1742-7843.2009.00518.x>.
- Mouly S, Roustit M, Bagheri H, Perault-Pochat M-C, Molimard M, Bordet R, et al. The French Levothyrox<sup>®</sup> crisis: we did the best we could but .... *Therapie*. 2019;74(3):431–5. <https://doi.org/10.1016/j.therap.2019.01.003>.