

Short Note

Not peer-reviewed version

Guidance on Multiplicity Analysis in Single-Trial Assessments: A No-Solution Equation

[Mondher Toumi](#)^{*}, [Bruno Fallissard](#), [Asma Jouini](#), Pascal Auquier, [Claude Dussart](#), [Laurent Boyer](#)

Posted Date: 29 April 2025

doi: 10.20944/preprints202504.2463.v1

Keywords: EU HTA; G-BA; HAS; HTA; HTACG; HTD; ICEMAN; JCA; JSC; MHT; MS; PICO



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Editorial

Guidance on Multiplicity Analysis in Single-Trial Assessments: A No-Solution Equation

Mondher Toumi ^{1,*}, Bruno Fallissard ², Asma Jouini ³, Pascal Auquier ¹, Claude Dussart ^{4,5} and Laurent Boyer ¹

¹ CEReSS/UR3279—Health Services Research and Quality of Life Center, Aix-Marseille University, 13385 Marseille, France

² CESP, INSERM U1018, Université Paris-Saclay, Villejuif, France

³ Clever Access, 1058, Tunis, Tunisia

⁴ UR 4129 P2S Parcours Santé Systémique, Université Claude Bernard Lyon 1, 7-11 rue Guillaume Paradin, Lyon, 69008, France

⁵ Pharmacie Centrale, Hospices Civils de Lyon, 57 Rue Francisque Darcieux, Saint-Genis-Laval, Lyon, 69230, France

* Correspondence: mto@inovintell.com

The European Health Technology Assessment (EU HTA) guidance on Multiplicity of Hypothesis Testing (MHT) [1] resembles an algebra no-solution equation. The hypothetico-deductive method used in inferential statistics, which this guidance references, relies on conclusive reasoning. However, the guidance fails to provide a clear resolution, leaving the matter to individual Member States (MS) despite its critical impact on statistical precision, a key pillar of validity in clinical trials for comparative effectiveness assessments in Joint Clinical Assessments (JCAs) [2].

MHT in inferential statistics increases the likelihood of false-positive results (Type I errors) [3]. Two main approaches exist for addressing this issue: (1) mathematical models such as the Bonferroni correction, which inflates sample size, and (2) hierarchical hypothesis testing, which does not [3,4]. While the guidance acknowledges the problem, it neither resolves nor mitigates it.

Key unresolved questions include:

- Should each PICO (Population, Intervention, Comparison, Outcome) be considered an independent predefined analysis, thereby exempting it from MHT? If so, it must be defined before results are available, which is currently not the case.
- Why should individual MS determine how to handle MHT, despite its status as a state-of-the-art statistical method? The Member State Coordination Group on Health Technology Assessment (HTACG) deviates from the EU HTA Regulation, which mandates adherence to state of the art rigorous methodology.

Moreover, the guidance states that "*for some MSs, the unplanned post hoc analysis is more important for the PICO than the planned analysis...in such cases, JCA report will document the p-value, and it will be marked as nominal*" [1]. This contradicts best scientific practices, where post hoc analysis p-values should not be reported since their statistical assumptions do not hold. Instead, such analyses should be classified as descriptive exploratory findings without statistical testing [5]. Requesting multiple PICOs during the scoping phase, once clinical trial results are available and submitted to the HTACG, means that any non-predefined PICO analysis qualifies as post hoc and should not be subjected to hypothesis testing. If testing is conducted, MHT adjustment must be applied, but p-values should not be reported to avoid misleading conclusions.

The guidance requires identifying whether post hoc analyses originate from authorities or sponsors [1]. However, given that the JCA subgroup is "decontextualized and judgment-free," [6] post hoc analyses should be treated equally, regardless of the requester. Identifying the originator

introduces implicit judgment, contradicting this principle. This contradiction invites a critical question: is the JCA process genuinely decontextualized, or is this principle selectively applied?

Moreover, this guidance clearly violates established principles of hypothetico-deductive reasoning and MHT by allowing assessors to arbitrarily select a single time point, preferably the last data cut, in multiple time point analyses. This introduces subjective judgment, which does not align with the EU HTA regulation's aim for objectivity and judgment-free JCAs [6]. Furthermore, the appropriateness of the chosen data cut-off depends on the endpoint and may affect the accuracy of the results: while it may be reasonable for hazard ratios or median survival, it is not for probability of survival (e.g., a time-dependent measure that requires a complete view of the survival curve) [7].

The guidance specifies that if a MS requested a specific subgroup analysis, the results must be reported and the Credibility of Effect Modification Analyses (ICEMAN) [8] criteria may be used to interpret and assess the results of submitted subgroup analyses. However, the ICEMAN checklist has been criticized for not considering the MHT related risk and its subjective judgement rating on a Visual Analogic Scale [9]. Again, free judgement is violated.

The introduction of additional post hoc analyses and the use of the ICEMAN tool for effect modifiers are useful for HTA appraisals but not for a decontextualized clinical assessment. However, their inclusion contradicts the EU HTA regulation, which mandates adherence to state-of-the-art methodology, while ICEMAN is instead a heuristic approach used to assess credibility rather than ensure statistical rigor.

The guidance fails to address how the multiplicity of hypothesis testing—both within a single PICO and across multiple PICOs—impacts the degree of certainty of the comparative effectiveness of an assessed intervention versus a reference. This is the primary goal of the JCA subgroup [2].

In conclusion, the guidance provides a detailed discussion of multiplicity but fails to align with contemporary well established statistical standards. The post hoc nature of PICOs defined after trial results renders them unsuitable for hypothesis testing; therefore, no statistical tests should be performed, and p-values should not be reported to be aligned with EU HTA Regulation.

The lack of systematic literature review, transparency, and accountability prevented the guidance's authors from identifying viable solutions to these critical methodological issues. It remains unclear whether the guidance considers the implications of MHT on statistical precision and certainty in comparative effectiveness assessments, the goal of JCA.

This ambiguity likely arises from diverging perspectives between HTA bodies—HAS (National French Health Authority) [10], which demands strict control over multiple hypotheses testing, and G-BA (The German Federal Joint Committee) which adopts a more flexible approach [11,12]. Differences stem from Fisher's and Neyman-Pearson's approaches to hypothesis testing. Fisher used it for inductive inference, where the p-value assesses data against the null hypothesis without considering alternatives. Conversely, Neyman and Pearson viewed hypothesis testing as "inductive behavior," making decisions (accepting or rejecting an alternative hypothesis) based on error probabilities rather than evidence quantification [13–15].

The HTACG faces two choices: either exclude MHT from its guidance, failing to address a key issue, or include it without resolving its complexities.

Given the constraints of the EU HTA Regulation, which limits judgment and flexibility while mandating evidence for multiple PICOs [6] (up to 13 in recent simulations [16]), resolving the MHT issue within the current framework is infeasible. The HTACG should withdraw this guidance and leave the matter to national authorities.

Author Contributions: M. T.: Conceptualized the content and wrote the first draft of the manuscript. The co-authors: Challenged the concept, edited the manuscript, and refined arguments for clarity and coherence. All authors reviewed and approved the final version for submission.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: None.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EU	European Union
EU HTA	European Health Technology Assessment
G-BA	The German Federal Joint Committee
HAS	National French Health Authority
HTA	Health Technology Assessment
HTACG	The Member State Coordination Group on Health Technology Assessment
HTD	Health Technology Developers
ICEMAN	Instrument to assess the Credibility of Effect Modification Analyses
JCA	Joint clinical assessment
JSC	Joint Scientific Consultation
MHT	Multiplicity of Hypothesis Testing
MS	Member States
PICO	Population, Intervention, Comparison, Outcome

References

1. HTA Coordination Group (HTACG). Guidance on reporting requirements for multiplicity issues and subgroup, sensitivity and post hoc analyses in joint clinical assessments. Available online: https://health.ec.europa.eu/document/download/f2f00444-2427-4db9-8370-d984b7148653_en?filename=hta_multiplicity_jca_guidance_en.pdf (accessed on 8 January 2025).
2. HTA Coordination Group (HTACG). Guidance on the validity of clinical studies for joint clinical assessments. V1.0. Available online: https://health.ec.europa.eu/document/download/9f9dbfe4-078b-4959-9a07-df9167258772_en?filename=hta_clinical-studies-validity_guidance_en.pdf (accessed on 30 December 2024).
3. Dmitrienko, A. and R. D'Agostino, Sr. Traditional multiplicity adjustment methods in clinical trials. *Stat Med* **2013**, *32*, 5172-218.
4. The European Medicines Agency (EMA). Guideline on multiplicity issues in clinical trial - Draft. Available online: https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-multiplicity-issues-clinical-trials_en.pdf (accessed on 24 February 2025).
5. Heckman, M.G., J.M. Davis, 3rd, and C.S. Crowson. Post Hoc Power Calculations: An Inappropriate Method for Interpreting the Findings of a Research Study. *J Rheumatol* **2022**, *49*, 867-870.
6. European Commission. Regulation (EU) 2021/2282 of the European Parliament and of the Council of 15 December 2021 on health technology assessment and amending Directive 2011/24/EU. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2282> (accessed on 18 December 2024).
7. McCaw, Z.R., D.H. Kim, and L.J. Wei. Pitfall in the Design and Analysis of Comparative Oncology Trials With a Time-to-Event Endpoint and Recommendations. *JNCI Cancer Spectr* **2022**, *6*.
8. Schandelmaier, S., M. Briel, R. Varadhan, C.H. Schmid, N. Devasenapathy, R.A. Hayward, J. Gagnier, M. Borenstein, G. van der Heijden, I.J. Dahabreh, X. Sun, W. Sauerbrei, M. Walsh, J.P.A. Ioannidis, L. Thabane, and G.H. Guyatt. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *Cmaj* **2020**, *192*, E901-e906.
9. Liu, M., Y. Gao, L. Zheng, and J. Tian. Set reliable hypotheses when using ICEMAN to assess the credibility of subgroup analysis. *Intensive Care Med* **2024**, *50*, 2223-2224.

10. French National Authority for Health (HAS). Doctrine of the Commission de la Transparence (CT). Available online: https://www.has-sante.fr/upload/docs/application/pdf/2021-03/doctrine_ct.pdf (accessed on 26 February 2025).
11. The independent Institute for Quality and Efficiency in Health Care (IQWiG). General Methods version 8.0. Available online: https://www.iqwig.de/methoden/allgemeine-methoden_entwurf-fuer-version-8-0.pdf (accessed on 26 February 2025).
12. The German Federal Joint Committee (G-BA). Rules of Procedure of the Federal Joint Committee. Available online: <https://www.g-ba.de/richtlinien/42/> (accessed on 26 February 2025).
13. Neyman, J., *Silver jubilee of my dispute with Fisher*. **1961**.
14. Fisher, R.A. Scientific thought and the refinement of human reasoning. *Journal of the Operations Research Society of Japan* **1960**, 3, 1-10.
15. Lehmann, E.L. The Fisher, Neyman-Pearson theories of testing hypotheses: one theory or two? *Journal of the American statistical Association* **1993**, 88, 1242-1249.
16. HTA Coordination Group (HTACG). PICO exercises. Available online: https://health.ec.europa.eu/publications/pico-exercises_en (accessed on 26 February 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.