

A Critical Appraisal of the HTA CG's Guidance on Indirect Treatment Comparisons for Scenarios where the Desired Comparator Must be Estimated Using RWD Due to Lack of RCT Evidence

Volker Schuster*

SAI MedPartners, Switzerland

***Corresponding author:** Volker Schuster, SAI MedPartners, Neuhofstrasse 12, 6341 Baar, Switzerland

ARTICLE INFO

Received: 📅 August 12, 2024

Published: 📅 August 21, 2024

Citation: Volker Schuster. A Critical Appraisal of the HTA CG's Guidance on Indirect Treatment Comparisons for Scenarios where the Desired Comparator Must be Estimated Using RWD Due to Lack of RCT Evidence. *Biomed J Sci & Tech Res* 58(2)-2024. BJSTR. MS.ID.009135.

ABSTRACT

The European Union's Regulation 2021/2282 on Health Technology Assessment (HTA) has introduced a standardized approach to evaluating healthcare interventions through the Joint Clinical Assessment (JCA). The Practical Guideline for Quantitative Evidence Synthesis, adopted on March 8, 2024, provides essential guidance for conducting and evaluating direct and indirect treatment comparisons, emphasizing systematic reviews and the PICO (Population, Intervention, Comparator, Outcome) framework. Oncology patients and to a lesser extent rare disease patients are the first to be affected by this guideline. The aim of his article is to summarize and critically review the guidance on indirect treatment comparison methodology and analyze the challenges and opportunities for manufacturers of (ultra-) rare disease therapeutics including rare cancers as well as precision medicine pioneers. Precision medicine therapies e.g. biomarker-based precision medicine is not at all uncommon in oncology. In oncology, small trial populations and high unmet need make it sometimes unethical as well as scientifically questionable to go beyond a single arm trial. Hence, in practice the technical guidance on how to use indirect treatment comparisons and external control groups is crucial.

Keywords: Joint Clinical Assessment (JCA); European Union Health Technology Assessment (EU HTA); Indirect Treatment Comparison (ITC); Real-World-Evidence (RWE); Real-World-Data (RWD); External Control Arms; External Controls; Propensity Score Matching (PSM); Inverse Probability of Treatment Weighting (IPTW); Matching-Adjusted Indirect Comparison (MAIC); Single Arm Trial; Precision Medicine; Oncology

Abbreviations: JCA: Joint Clinical Assessment; EU HTA: European Union Health Technology Assessment; ITC: Indirect Treatment Comparison; RWE: Real-World-Evidence; RWD: Real-World-Data; PSM: Propensity Score Matching; IPTW: Inverse Probability of Treatment Weighting; MAIC: Matching-Adjusted Indirect Comparison; ITCs: Indirect Treatment Comparisons

Introduction

The European Union's Regulation 2021/2282 on Health Technology Assessment (HTA) has ushered in a transformative approach to evaluating healthcare interventions across member states [1,2]. At the heart of this process lies the Joint Clinical Assessment (JCA), a collaborative effort that demands rigorous, transparent, and harmonized methodologies [3-5]. The HTA Coordinating Group's Practical Guideline for Quantitative Evidence Synthesis, adopted on March 8, 2024, is intended to serve as a crucial roadmap for navigating this new landscape. It provides detailed guidance on conducting and eval-

uating direct and indirect treatment comparisons, ensuring that evidence synthesis processes meet the high standards required for EU-wide assessments [1]. For pharmaceutical companies, the guideline's emphasis on systematic and transparent approaches to evidence synthesis is paramount. The document stresses the importance of rigorous systematic reviews and careful study selection based on the PICO (Population, Intervention, Comparator, Outcome) framework. Central to the guideline is the concept of exchangeability, which underpins the validity of meta-analyses and network meta-analyses. This principle is operationalized through three key components:

1. **Similarity:** Evaluating study and patient characteristics, intervention and comparator characteristics, and outcome definitions to identify potential effect modifiers.
2. **Homogeneity:** Assessing the consistency of treatment effects across studies for each pairwise comparison.
3. **Consistency:** For indirect comparisons, ensuring that direct and indirect evidence are in agreement.

For the pharmaceutical industry, this emphasis on exchangeability necessitates a thorough understanding of their clinical trial data and how it compares to other relevant studies in the field. Companies will need to be prepared to demonstrate the comparability of their evidence with existing data, potentially influencing trial design and data collection strategies from the early stages of drug development [5]. The guideline delves into various methodological approaches for direct and indirect comparisons, including standard meta-analysis techniques, the Knapp-Hartung method for small sample sizes, and sophisticated network meta-analysis methods. Special attention is given to time-to-event data, a crucial consideration for many oncology and chronic disease treatments. This comprehensive coverage provides clarity on the analytical methods that will be accepted in JCAs, allowing pharmaceutical companies to align their evidence synthesis strategies accordingly. Recognizing the evolving nature of clinical evidence, the guideline also addresses population-adjusted methods for indirect comparisons and the use of non-randomized evidence. This inclusion is particularly relevant for the pharmaceutical industry, as it provides a framework for leveraging real-world evidence and addressing scenarios where traditional randomized controlled trials may be limited or unfeasible.

Throughout, the guideline emphasizes the importance of transparency and comprehensive reporting. It provides detailed requirements for what should be included in JCA reports, ensuring that member states have the information they need to make informed decisions.

Indirect Treatment Comparisons (ITCs) using Real-World Data (RWD) Evidence in Scenarios where the Desired Comparator must be Estimated Using RWD Due to Lack of RCT Evidence

The guideline addresses the use of non-randomized evidence in Section 6, "Assessment of comparisons based upon non-randomised evidence". Non-randomized studies are susceptible to confounding and selection bias. The guideline recommends using propensity score methods to address these biases. Assessors should critically evaluate the plausibility of the "no unmeasured confounders" assumption.

Propensity Scores (Section 6.2)

The guideline provides advice on checking the validity of propensity score matching and/or weighting assumptions. It emphasizes the importance of assessing covariate balance and overlap in propensity

score distributions. Interpretation of results should consider the population to which the treatment effect applies (e.g., average treatment effect vs. average treatment effect among the treated). The guideline addresses therapies for which the desired comparator cannot be estimated with RCT evidence primarily in Section 5, "Assessment of population-adjusted methods," particularly in the context of unanchored comparisons.

Unanchored MAICs and STCs (Section 5.5)

The guideline acknowledges the challenges of unanchored comparisons, where there is no common comparator between studies. It emphasizes that these methods rely on even stronger assumptions than anchored comparisons. The "no unmeasured effect modifiers" assumption is crucial and should be carefully evaluated.

Interpretation and Use of Population-Adjusted Results (Section 5.6)

Results from unanchored comparisons should be interpreted with extreme caution. The guideline suggests that such comparisons may be useful for hypothesis generation but are generally not suitable for decision-making. The guideline's approach to RWD and unanchored comparisons generally aligns with international standards, emphasizing caution and transparency [6,7]. The focus on propensity score methods and the assessment of their assumptions is consistent with best practices [8-10].

Discussion

While the guideline acknowledges the challenges of unanchored comparisons, it could provide more specific guidance on how to proceed when RCT evidence is unavailable for a desired comparator. More detailed advice on how to assess the plausibility of the "no unmeasured confounders" assumption would be beneficial. The guideline could offer more concrete recommendations on how to integrate RWD-derived comparisons into the overall assessment when RCT evidence is lacking. The guideline could provide a framework for integrating RWD-derived comparisons within a hierarchical evidence synthesis approach. This would involve explicitly stating how to weigh and combine evidence from different study designs, including RCTs and observational studies [11]. Moreover, the guideline could be more specific about bias-adjusted analysis for RWD-derived comparisons by recommending particular techniques like Propensity Score Matching (PSM) and Inverse Probability of Treatment Weighting (IPTW) [12,13]. It could provide step-by-step guidance on implementing these methods, including how to select covariates, assess balance, and conduct sensitivity analyses [14-17]. Additionally, the guideline could outline specific reporting requirements for these analyses, ensuring transparency and reproducibility. By providing more detailed guidance on PSM and IPTW, the guideline could enhance the quality and consistency of RWD-derived comparisons in HTA submissions.

More detailed guidance on conducting and interpreting quantitative bias analysis for RWD-derived comparisons could be provided, helping assessors to understand the potential impact of unmeasured confounding [18]. Specific Bayesian approaches for combining RCT and RWD evidence, which allow for the formal incorporation of external information and expert opinion, could be recommended [19]. More specific guidance on using target trial emulation techniques to design and analyze observational studies that mimic RCTs [20] could be provided. By incorporating these elements, the guideline could provide more concrete and comprehensive recommendations on integrating RWD-derived comparisons into the overall assessment when RCT evidence is lacking. This would align with international efforts to develop robust methodologies for incorporating real-world evidence into healthcare decision-making. Finally the guideline's basic interpretation of the RCT as the "gold standard" of efficacy and safety evidence [1] is underwhelming as it does not fully acknowledge the inherent requirements of ultra rare diseases and above all precision medicine. If we take precision medicine seriously, regulatory as well as HTA procedures need to be mindful of the fact, that having an individual treatment for each individual tumor and patient limits the benefits of randomized comparisons.

Limited patient cohorts that require regulatory and HTA decisions based on single arm trial data need hence be seen as an opportunity to secure knowledge, gather experience and build scientific best practice consensus. This consensus should include the general requirements to put on real-world-data to be eligible to be utilized as an external control arm to a clinical trial and it should certainly point out as well what specific methods should and should not be utilized in deriving external controls from such real-world-data sources. To be fully prepared for precision medicine, we need to be able to find reliable and reproducible methods as well as clear scientific consensus on the if and the how to develop "synthetic external controls" from one or multiple real-world data sources [21-23]. Once the promise of precision medicine fully comes true and treatments get fully individualized (not just tailored to a cohort), there will be no patients to randomize any longer unless we are methodologically fully equipped to develop a synthetic counterpart for each patient. AI and advanced biotech will likely enable us to tailor therapies to individual tumors and patients in the very near future. It seems highly plausible, that developing clear and comprehensive guidelines to generate or derive "synthetic" patient counterparts prospectively for each individual patient from multiple real-world-data sources through generative AI might speed up clinical development while still capturing the true characteristics of real-life patients in contrast to the often idealized selection of clinical trial patients. Rather than rejecting real-world-data sources for their shortcomings and lack of documentation, we should demand usage of all available real-world-data and develop clear and transparent standards for AI to acknowledge and account for all of the shortcomings in a reasonable, transparent and reproducible way. RCTs can be

only considered a "gold standard", if internal validity is valued more than external validity. If the scientific consensus would be to use all real-world data that meet predefined criteria, this consensus would rule out all possibilities of cherry picking. A clear set of reproducible rules, that guide a prospective AI-driven generation of "synthetic" external controls might become as reliable as randomization. Whoever wants to reduce doubt on the potential of real-world-evidence, needs to work towards universally agreed-upon standards of generating and curating real-world-data by independent and trustworthy agents like medical associations or patient organizations.

In conclusion, while the guideline provides a solid foundation for assessing ITCs using RWD and unanchored comparisons, there's room for more specific guidance on handling scenarios where RCT evidence is unavailable for key comparators. This is particularly important as this guideline is going to affect cancer patients and first.

References

1. (2024) HTA CG Methodological Guideline for Quantitative Evidence Synthesis: Direct and Indirect Comparisons. HTA CG 3(7).
2. (2021) 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. Official Journal of the European Union L 458(1): 1-32.
3. (2024) The EU regulation on health technology assessment: what's in it and why it matters? European Commission.
4. (2024) Document 32024R1381, Commission Implementing Regulation (EU) 2024/1381 of 23 May 2024 laying down, pursuant to Regulation (EU) 2021 EUR Lex.
5. Schuster V (2024) EU HTA Regulation and Joint Clinical Assessment-Threat or Opportunity? J Market Access Health Policy 12(2): 100-104.
6. Berger M L, Sox H, Willke R J, Brixner D L, Eichler H G, et al. (2017) Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiology and drug safety* 26(9): 1033-1039.
7. Franklin J M, Pawar A, Martin D, Glynn R J, Levenson M, et al. (2020) Non-randomized real-world evidence to support regulatory decision making: process for a randomized trial replication project. *Clinical Pharmacology & Therapeutics* 107(4): 817-826.
8. Austin P C (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research* 46(3): 399-424.
9. Rosenbaum P R, Rubin D B (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41-55.
10. Stuart E A (2010) Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics* 25(1): 1-21.
11. Efthimiou O, Mavridis D, Debray T P, Samara M, Belger M, et al. (2017) Combining randomized and non-randomized evidence in network meta-analysis. *Statistics in medicine* 36(8): 1210-1226.
12. Rosenbaum P R, Rubin D B (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41-55.

13. Austin P C (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research* 46(3): 399-424.
14. Stuart E A (2010) Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics* 25(1): 1-21.
15. Imbens G W, Rubin D B (2015) *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
16. Hernán M A, Robins J M (2020) *Causal inference: what if*. Boca Raton Chapman & Hall/CRC, pp. 1-359.
17. Turner R M, Spiegelhalter D J, Smith G C S, Thompson S G (2009) Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172(1): 21-47.
18. Stürmer, Schneeweiss, Avorn, Glynn R J (2005) Adjusting effect estimates for unmeasured confounding with validation data using propensity score calibration. *American Journal of Epidemiology* 162(3): 279-289.
19. Lash T L, Fox M P, MacLehose R, McCandless L C, Greenland S, et al. (2014) Good practices for quantitative bias analysis. *International journal of epidemiology*, 43(6): 1969-1985.
20. Spiegelhalter D J, Abrams K R, J P Myles (2004) *Bayesian approaches to clinical trials and health-care evaluation*. John Wiley & Sons (13).
21. Wang SV, Schneeweiss S, Franklin JM, Desai RJ, Feldman W, et al. (2023) Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA* 329(16): 1376-1385.
22. Struebing A, McKibbin C, Ruan H, Mackay E, Dennis N, et al. (2024) Augmenting external control arms using Bayesian borrowing: a case study in first-line non-small cell lung cancer. *J Comp Eff Res* 13(5): e230175.
23. Hernán MA, Robins JM (2016) Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 183(8): 758-764.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2024.58.009135

Volker Schuster. Biomed J Sci & Tech Res



This work is licensed under Creative Commons Attribution 4.0 License

Submission Link: <https://biomedres.us/submit-manuscript.php>



Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<https://biomedres.us/>