ECKER ECKER

P041

Beuermann J., Grafe V., Rmo M., Oldeland J. | Ecker + Ecker GmbH | Hamburg | Germany

Absolute standardized Differences – an objective tool for checking similarity in indirect treatment comparisons?

Background

Indirect treatment comparisons (ITCs) are an important tool in HTA decision-making. A key assumption is exchangeability, requiring similarity, homogeneity, and consistency between studies. Similarity is commonly assessed by comparing baseline distributions of effect modifiers and prognostic factors, although this is often done visually rather than quantitatively. We discuss the use of absolute standardized differences and the derivation of an overall similarity threshold.

Following approaches used to assess balance in propensity score methods, we propose the use of absolute standardized differences to evaluate similarity in baseline characteristics. While thresholds for balance assessment in propensity score analyses (e.g., 0.1–0.25) are suggested in the EU HTA guidance on evidence synthesis [1], no established thresholds currently exist for indirect comparisons based on randomized evidence in HTA.

Methods

To derive a possible threshold value, we evaluated all HTA procedures carried out by the German HTA body, the G-BA, from 2021 to 2025 [2]. For ITCs deemed methodologically sound and accepted by the G-BA, we extracted the reported baseline characteristics (continuous and categorical) from the studies employed for the indirect comparison and calculated the absolute standardized differences.

Results

We identified 11 approved ITCs comprising 97 evaluable baseline characteristics. All ITCs applied the Bucher method. Absolute standardized differences (ASDs) were aggregated, and four potential similarity thresholds with decreasing restrictiveness were derived from their distribution (10th percentile, median, mean, and 90th percentile; *Fig. 1*). A median ASD of 0.18 IQR [0.10, 0.43] was observed, ASDs for continuous baseline characteristics were slightly lower than those for categorical characteristics (*Fig. 2*). No substantial differences were observed between oncological and non-oncological indications (*Fig. 2*). Among baseline characteristics evaluable in at least five ITCs, region showed the highest average dissimilarity (*Fig. 3*).

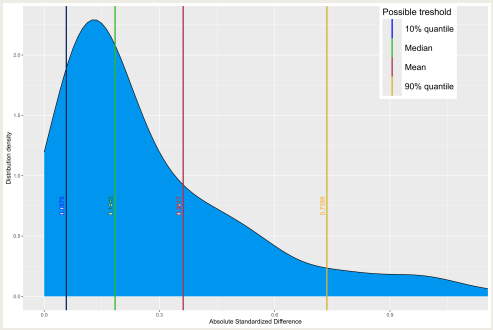

Fig. 1: Distribution of ASDs across accepted ITCs including possible tresholds

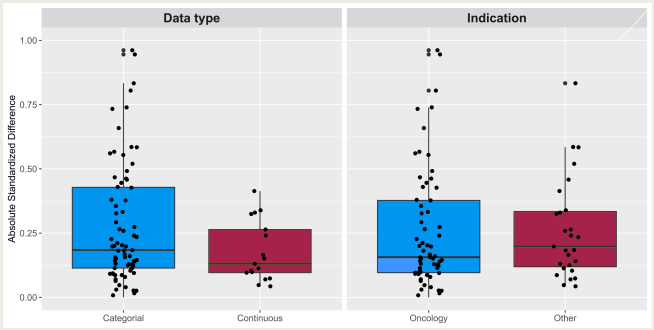

Fig. 2: Distribution of ASD by variable type and indication

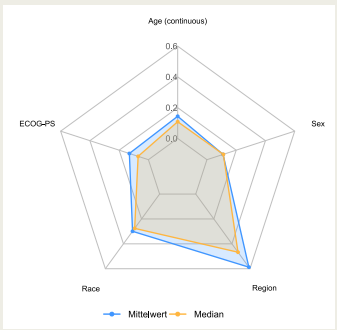

Fig. 3: Distribution of ASD by baseline characteristics

Limitations

- Limited ability to calculate comprehensive ASDs due to unreported baseline characteristics.
- Although all evaluable baseline characteristics were included, assessment of the similarity assumption primarily depends on the availability and comparability of effect modifiers and prognostic factors.

Conclusion

Assessing similarity remains context-dependent and complex; however, ASD calculations can provide an initial quantitative overview. Based on the observed ASD distribution, a threshold of 0.2 appears to be a pragmatic and moderately restrictive option. This corresponds to the rounded median ASD and lies within the commonly used propensity score balance range of 0.1–0.25. Given the right-skewed ASD distribution, the median was considered more robust than the mean for deriving a practical threshold.

References

[1] EU HTA Practical Guideline, www.health.ec.europa.eu

[2] Gemeinsamer Bundesausschuss (G-BA), www.g-ba.de

SCAN ME