FUNDED BY NIHR | National Institute for Health and Care Research

HTA130

Evaluating Clinical Similarity: a Systematic Review of Methods Used Within National Institute for Health and Care Excellence (NICE) Cost-Comparison Evaluations (CCEs) When Relying on Statistically Non-Significant Differences from Indirect Treatment Comparisons (ITCs)

Nicole Downes, BSc (Hons); Archie Walters, MSc; Ben Burgess, PhD; Sophie Ip, PhD; Clare Dadswell, PhD; Steven J Edwards, DPhil
BMJ Technology Assessment Group (BMJ-TAG), BMJ Group, London, United Kingdom.

Introduction

- CCEs were introduced by NICE in 2017 as a lower-resource alternative to the Single Technology Appraisal process.[1-3]
- However, validating the assumption that new treatments are clinically no worse than relevant comparator treatments is challenging, particularly when these assumptions are based on statistically non-significant differences from ITCs.[1, 3]
- While indicating that there is insufficient evidence to demonstrate a statistically significant difference between treatments, non-significant differences cannot prove similarity, non-inferiority or equivalence.[4]
- As there is no guidance on how to interpret such results to support or refute conclusions of clinical similarity, there is potential for inconsistency in interpretation across NICE CCEs, which may impact decision-making.

Objective

The aim of this research was to review and summarise the basis for concluding clinical similarity in NICE CCEs that rely on statistically non-significant results from ITCs, including current methods used to interpret the ITC results and whether there are any inconsistencies in decision-making within these CCEs.

BMJ Technology Assessment Group (BMJ-TAG), part of BMJ Group, conducts independent health research, utilising the expertise of its teams of clinical evidence analysts and health economists

Methods

- After searches of the NICE website, review of appraisals included in similar reviews and contact with NICE in April 2025, NICE CCEs from 2017 onwards were included if they relied on statistically non-significant ITC results for at least one comparator and had final guidance published.
- Details on the methods used to support decision-making and associated rationale were extracted, with narrative syntheses of the company, EAG and committee perspectives across appraisals performed.

Results

In total, 41 NICE CCEs were included. Over 60% used network meta-analyses (NMAs) as the ITC method. A summary of company, EAG and committee perspectives is provided in Table 1. Key findings were:

- EAGs frequently raised concerns about relying on statistically non-significant ITC results but companies often strongly emphasised these results;
- Formal methods to explore uncertainty in ITC results were used in a minority (~32.0%) of appraisals by the EAG and/or company. This included probabilities from NMAs such as the probability of being the best treatment or being non-inferior, and using thresholds such as minimal clinically important differences (MCIDs) or non-inferiority margins (NIMs) to support interpretation of the point estimate and confidence intervals;
- No insights into committee preferences for the interpretation of such ITC results were available, but supplementary sources of evidence, such as clinical expert feedback, often reduced uncertainty.

Table 1. Summary of company, EAG and committee methods and perspectives

Company	EAG	Committee
• Emphasis on the lack of statistically significant differences with point estimate size and direction considered; • Point estimates favouring the intervention sometimes interpreted as evidence that clinical similarity was conservative; • ~15% mentioned probabilities from NMAs and ~10% used MCIDs or NIMs. In some cases, this was only after an EAG request.	• Highlighted limitations of statistically non-significant results; • Supportive evidence such as alternative analyses or clinical expert feedback not always sufficient to reduce EAG concerns; • NMA probabilities or clinical thresholds mentioned or used more often by EAGs than companies (~17% each).	• No clear position on the interpretation of statistically non-significant ITC results was identified; • Supportive evidence such as information on the mechanism of action and clinical expert feedback, often reduced uncertainty in decision-making.

Abbreviations: EAG, External Assessment Group; ITC, indirect treatment comparison; MCID, minimal clinically important difference; NIM, non-inferiority margin; NMA, network meta-analysis.

Conclusions

- Inconsistencies in the interpretation of statistically non-significant ITC results exist within individual NICE CCEs and across different appraisals.
- Formal methods to explore uncertainty in these ITCs are only being used in a minority of these appraisals.
- Wider application of such methods in CCEs relying on non-significant ITC results may improve consistency across CCEs, which may be facilitated by additional guidance from NICE for companies and EAGs.

References

1. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual, 2023. Available from: https://www.nice.org.uk/process/pmg36. Date accessed: Apr 25.

2. Thurgar E. The Evolution of Cost Comparison at NICE: What Can We Learn from Fast-Track Appraisals? 2022 https://www.ispor.org/ docs/default-source/euro2022/thurgar-e-1pdf.pdf?sfvrsn=494aa432 _0. ISPOR Europe; 6-9 November Vienna, Austria.

3. Armand J, Fontrier AM, Sawyer L. Cost-comparison, an easier route to NICE recommendations? 2023. p. https://www.ispor.org/docs/ default-source/euro2023/isporeurope23armandhta357poster132726-pdf.pdf?sfvrsn=12a50a55_0. ISPOR Europe; 12-15 November; Copenhagen, Denmark.

4. Gewandter JS, McDermott MP, Kitt RA, Chaudari J, Koch JG, Evans SR, *et al.* Interpretation of CIs in clinical trials with non-significant results: systematic review and recommendations. *BMJ Open* 2017; **7**: e017288.