Lumanity

Targeted review: generalizability of clinical trial evidence to the patient population in NICE technology appraisals

Authors: Wijenayake N[1], Thornton I[2], Smith M[1], Zhi Tan Y[3]

Affiliations: [1] Lumanity, London, UK; [2] Lumanity, Sheffield, UK; [3]Lumanity, Utrecht, The Netherlands

INTRODUCTION

- Determining the relevance and usability of evidence for evidence-based decision making is an essential component of health technology assessment (HTA). Globally, HTA bodies strive to base decisions on data relevant to our real-world populations[1]
- The European Regulation on HTA emphasizes terms like 'external validity', 'applicability' and 'transferability', as they relate to applying study results in real-world settings.[1] The UK's National Institute for Health and Care Excellence (NICE) health technology evaluations manual highlights the need for clinical trial evidence to reflect UK-specific patients, comparators, outcomes and resource use, but lacks a formal assessment framework, likely to allow a case-by-case review.
- International clinical trials often use selective patient groups based on factors like age, comorbidities, and treatment pathways, which may not align with UK National Health Service (NHS) practice. This complicates data application to specific populations, highlighting the need to ensure relevance to UK patients for informed decisions
- Cancer outcomes and services are a key NHS priority in the UK. There is active research in a range of women's cancers, where there is a steady increase in age-standardised incidence rates. [2]

OBJECTIVES

Our work examined how the generalizability of clinical trial data to UK populations is addressed across different stages of NICE's appraisal process. By reviewing recent technology appraisals (TAs), we explored explicit considerations of external validity from companies, External Assessment Groups (EAGs), and NICE, focusing on patient demographics, treatment comparators and clinical outcomes.

METHODS

- The review focused on five types of women's cancer: ovarian, cervical, endometrial, vaginal, and vulval. TAs with published Final Appraisal Determination (FAD) between 2019 and 2024 were identified from the NICE website
- Initial screening of TAs was conducted by one reviewer and cross-checked independently by another. Review and extraction processes were guided by the following search terms: "generalis", "generaliz", "validat", "UK population", "population in England", "prognostic", "effect modif", and "clinical practice". Data and relevant information were systematically extracted based on these pre-specified search terms
- Extractions were organized into the following categories, in the context of clinical trial or UK population generalizability: population demographics; baseline clinical parameters; clinical interventions and comparators (including subsequent therapies); and explicit comparability of trial and real-world outcomes
  - These aspects were identified using the population, intervention, comparator, outcome (PICO) framework to standardize key elements of the decision problem and systematically assess data relevance and applicability
- Data were extracted in chronological order from the Company Submission (CS), EAG reports, Public Committee Slides (PCS), and FAD. If multiple CSs of the same appraisal were available, the most recent version was extracted. The number of hits per category was recorded and analysed

RESULTS

- A total of seventeen TAs were identified, all in cervical, ovarian, or endometrial cancer (Table 1)
  - Of the seventeen TAs, data extraction was completed for ten[3–12] and seven were excluded due to being terminated, incomplete, or superseded
- Every TA considered at least one aspect of the four categories at the CS, EAG and PCS stage, with the primary focus on aligning population demographics and clinical trial interventions and comparators with UK clinical practice (mentioned in 90% of all reviewed documents)
- Clear trends emerged across different stages of the TA process, showing a decline in mentions of all categories as the appraisal process progressed (Figure 1). Of the categories mentioned, issues were deemed an area of uncertainty that required resolving (Figure 2) or not (Figure 3)

Table 1. Summary of the 10 TAs included in the data extraction

Publication date	Disease area	Intervention	Reference
17/02/2021	Ovarian cancer	Niraparib	TA673[3]
16/03/2022	Endometrial cancer	Dostarlimab	TA779[4]
20/04/2022	Ovarian cancer	Niraparib	TA784[5]
17/01/2023	Ovarian cancer	Olaparib with bevacizumab	TA946[6]
28/03/2024	Ovarian cancer	Olaparib	TA962[7]
3/04/2024	Endometrial cancer	Dostarlimab	TA963[8]
21/06/2023	Endometrial cancer	Pembrolizumab	TA904[9]
5/07/2023	Ovarian cancer	Olaparib	TA908[10]
20/09/2023	Endometrial cancer	Pembrolizumab	TA914[11]
13/12/2023	Cervical cancer	Pembrolizumab plus chemotherapy with or without bevacizumab	TA939[12]

Company submission stage

- Companies explicitly commented on at least one of the four categories in 100% of TAs. In 70% of mentions, the company's conclusion was that the evidence was appropriate for UK decision making
- On the occasions that the company's conclusion was that the evidence was not appropriate for UK decision making, 50% of the uncertainty this was a consideration of whether trial-based subsequent treatments aligned with standard UK practice

EAG stage

- Most EAG reports (60%) commented on more aspects of external validity than company submissions, 30% addressed the same number, and 10% addressed fewer
- When comparing categories of external validity explored by both the CS and EAG, the level of agreement in the conclusion as to whether it was considered an issue of generalizability or not was approximately 45%
  - Among matched topics, the EAG frequently questioned the comparability of trial population demographics and clinical baseline characteristics to UK practice. Concerns often focused on whether the observed treatment effect might be larger than in real-world settings, due to differences in patient characteristics
  - Resolutions to these issues, which were explicitly stated in five TAs, involved requests for additional clinical trial or real-world data, clinical validation, or scenario analyses

Table 2. Number of total mentions by category across all TAs, at each stage of the NICE appraisal process

Category	CS	EAG	PCS	FAD
Population demographic	8	8	7	7
Baseline clinical parameters	2	1	0	1
Clinical trial interventions and comparators	7	9	5	4
Clinical outcomes in comparator arm	2	2	2	0
Other	1	1	0	1
Total mentions	**20**	**21**	**14**	**13**

Public committee slides stage

- The public committee slides mentioned aspects of generalizability less frequently than EAG reports (14 vs 21 mentions)
  - Only three of four categories were covered at the PS stage, with 50% of mentions (n = 7[3-4,6-7,9,11-12]) regarding population demographics (primarily ethnicity and age differences from the UK population), 35% regarding clinical trial interventions and comparators including subsequent therapies, and 15% regarding comparator-arm outcomes including post-progression survival
  - Additional concerns included uncertainty associated with treatment sequencing and challenges associated with interpreting immature trial data

FAD stage

- The FAD results indicate that the relevant areas of uncertainty were largely resolved before publication and/or deemed to not be a major issue when considering the full appraisal to warrant further discussion

Figure 2. Number of total mentions of a category that was deemed an area of uncertainty that required resolving, at each stage of the NICE appraisal process

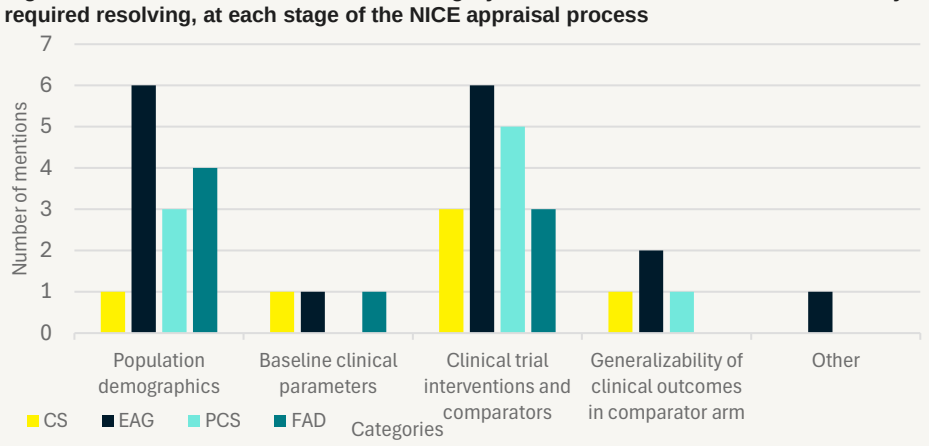
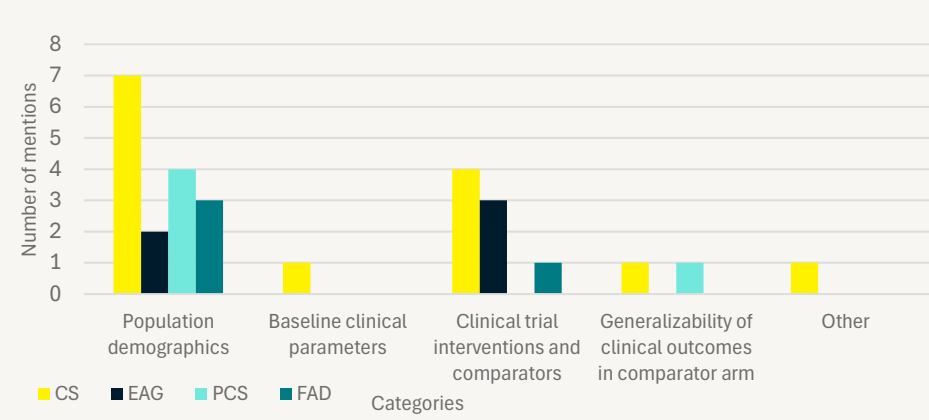


Figure 3. Number of total mentions of a category that was not deemed an area of uncertainty that required resolving, at each stage of the NICE appraisal process



CONCLUSIONS

- Ultimately, 100% of the reviewed appraisals received positive recommendations.
- This analysis highlighted that ensuring trial data aligns with UK clinical practice is a key focus in NICE technology appraisals, as shown by its frequent mention. However, results indicate varying emphasis across appraisals and by stage in the process
- The stages between the EAG report and the public committee meeting are crucial for resolving issues, reflected by the number of issues resolved following the EAG report stage
  - Companies often address queries on population generalizability through additional data or validation
  - Technical engagement, which played a key role in the NICE process before 2022, is now optional. This may impact future trends as the number of opportunities to discuss or address areas of uncertainty is reduced
- There remains relatively high levels of disagreement between companies and EAGs. Development of comprehensive, standardized criteria for reviewing the applicability of evidence for HTAs could support consistency and transparency in HTA decisions
- This analysis only considered the final guidance for each TA; future research could provide greater insight by reviewing draft guidance and document iterations

REFERENCES

1. HTA CG. 2024. Clinical Studies Validity Guidance. Accessed: 2 June 2024. 2. Sun et al. 2024. EClinicalMedicine, 74. 3. NICE. 2021. TA673. Accessed: 22 May 2024. 4. NICE. 2022. TA779. Accessed: 24 May 2024. 5. NICE. 2022. TA784. Accessed: 1 June 2024. 6. NICE. 2024. TA946. Accessed: 1 June 2024. 7. NICE. 2024. TA962. Accessed: 22 May 2024. 8. NICE. 2024. TA963. Accessed: 1 June 2024. 9. NICE. 2023. TA904. Accessed: 22 May 2024. 10. NICE. 2023. TA908. Accessed: 20 May 2024. 11. NICE. 2024. TA914. Accessed: 22 May 2024. 12. NICE. 2023. TA939. Accessed: 1 June 2024.

An electronic version of the poster can be viewed by scanning the QR code.

Poster presented at the 2024 ISPOR Europe; 17-20 November 2024; Barcelona, Spain.