Comparing ROBINS-I V2 and Downs & Black for nRCTs in the Context of HTA

HTA84

Yogesh Punekar[1], *Ranjith Babu Vagichalra*[2], *Aditi Agarwal*[2], *Manne Mithun Chakrawarthy*[2], *Premsagar Kummari*[2], *Veeranki Prasanth*[2] , *Geddamuri Bala Ganesh*[2]
[1]*IQVIA, London, United Kingdom*
[2]*IQVIA, India*

BACKGROUND AND OBJECTIVES

Health technology assessment (HTA) bodies are increasingly recognising the value of non-randomised clinical trials (nRCTs), particularly when randomised clinical trials (RCTs) data is limited or not available. Risk of bias in non-randomised studies of interventions (ROBINS-I) (1) and the Downs & Black checklists (2) are widely used critical appraisal tools for evaluating nRCTs. The Downs & Black tool, published in 1998, is valued for its simplicity and time efficiency. In contrast, the updated ROBINS-I Version 2 (V2), released in November 2024, offers enhanced detail, refined signalling questions, and improved judgement logic. This review compares ROBINS-I V2 and Downs & Black using published nRCTs, focusing on domain-level outputs and complexity in the context of past HTA recommendations.

METHODOLOGY

- Three published nRCTs in endometrial cancer were assessed using both appraisal tools
- Domain-level judgements (evaluated across 10 domains), scoring styles, and usability were compared between the two tools
- European Union (EU) HTA methods guidelines from National Institute for Health and Care Excellence (NICE), Scottish Medicines Consortium (SMC), National Centre for Pharmacoeconomics (NCPE), French National Authority for Health (HAS), and Federal Joint Committee (G-BA) were reviewed to understand current study assessment practices

RESULTS

- Three nRCTs studies were evaluated using both tools. ROBINS-I V2 identified moderate to serious bias, particularly from confounding and selective reporting with overall serious to critical in two studies despite low risk in other domains
- The same studies scored favourably (18-24/27) using the Downs & Black tool, highlighting strong reporting and external validity but offering limited details in bias-specific confounder, lost to follow-up, and power (Supplementary Table 1)
- Comparison of studies assessed using both tools, observed a key difference that ROBINS-I V2 employs a qualitative, algorithm-based approach requiring an expert judgment, whereas the Downs & Black tool uses a quantitative checklist method with less detailed scoring guidance
- A comparison of the two tools and the critical appraisal requirements for nRCTs in EU HTAs are summarised in Table 1 and Table 2 respectively

Table 1: Comparison of the ROBINS-I V2 with Downs & black checklist domains

Domain	ROBINS-I V2 (1)	ROBINS-I V2 Comments (1)	Downs & Black Checklist (2)	Downs & Black checklist Comments (2)
Bias due to confounding	✓	Q1.1–1.5 assess if baseline confounders were reported for (e.g., "Were there important confounding domains not considered or controlled for?").	✓*	Item 5: "Are distributions of principal confounders described?" Limited; no structured approach to confounder identification or adjustment is included.
Bias in classification of interventions	✓	Q2.1–2.5 address intervention classification and misclassification (e.g., "Was intervention status accurately classified for all or nearly all participants?").	✓	Item 4: "Were interventions clearly described?" Ensures clarity but does not explore misclassification risk systematically.
Bias in selection of participants	✓	Q3.1–3.10 evaluate if participant inclusion led to bias (e.g., "Were eligible participants representative of the population?").	✓	Item 21: "Were subjects representative of the source population?" Also reflects external validity rather than bias alone.
Bias due to deviations from intended interventions	✓	Q4.1–4.5 assess non-adherence, co-interventions, and whether deviations introduced bias (e.g., "Were deviations balanced or affected outcome?").	✓	Item 6: "Were the main findings of the study clearly described?" Simpler compliance check, lacks causal implications or assessment of deviations.
Bias due to missing data	✓	Q5.1–5.11 examine extent, reasons, and handling of missing data (e.g., "Is it likely that missing data could bias the results?").	✓	Items 9 and 26 cover reporting and justification of attrition and loss to follow-up (e.g., "Were loss to follow-up described or were loss to follow-up taken into account").
Bias in measurement of outcomes	✓	Q6.1–6.4 determine if outcome assessment was blinded or influenced by intervention knowledge (e.g., "Could outcome measurement be influenced by bias?").	✓	Items 7–8: "Were the outcome measures valid and reliable?" and "Were they applied equally to all subjects?" Relates to accuracy and consistency.
Bias in selection of reported result	✓	Q7.1–7.3 evaluate if prespecified outcomes were all reported (e.g., "Was there evidence of selective reporting of outcomes or timepoints?").	✓	Item 16: "Were all important outcomes considered in analysis?" Implies thoroughness but lacks structure to detect selective reporting.
External validity (generalizability)	×	Not assessed. ROBINS-I is designed to evaluate internal validity only.	✓	Items 11–13: "Were staff and representative members blinded?" Addresses applicability of findings to real-world populations.
Reporting quality (clarity)	×	Not covered in ROBINS-I (reporting clarity is not the same as bias).	✓	Items 1–3, 10, 17–20 assess clarity of objectives, methods, variability, and adverse event reporting. Evaluates completeness and transparency of reporting.
Power/sample size estimation	×	Not included. ROBINS-I does not assess adequacy of power or sample size.	✓	Item 27: "Did the study have sufficient power to detect a clinically important effect?" One item assesses design robustness.

Abbreviations: ROBINS-I V2, Risk of Bias in Non-randomized Studies of Interventions, Version 2, X- Absent, ✓ - Present, *: Partially addressed

Table 2: Critical appraisal tool requirements for nRCTs in EU HTAs

Submission requirements / HTA Agencies	United Kingdom NICE (3)	Scotland SMC*(4)	Ireland NCPE*(5)	France HAS*(6)	Germany G-BA(7)	European Union (EUnetHTA to be updated by JCA) (7)
Critical appraisal tool required	ROBINS-I	Any appropriate tool	Any appropriate tool	Any appropriate tool	ROBINS-I	ROBINS-I

*Any appropriate tool is recommended while submission
Abbreviations: EU HTA, European Union health technology assessment; G-BA, Federal Joint Committee; HAS, French National Authority for Health NCPE, National Centre for Pharmacoeconomics; NICE, National Institute for Health and Care Excellence; nRCT, Non-randomised trial; ROBINS-I V2, Risk of Bias in Non-randomized Studies of Interventions, Version 2; SMC, Scottish Medicines Consortium;

CONCLUSION

- ROBINS-I V2 provides a structured, bias-oriented assessment that captures domain-specific risks such as confounding and selective reporting. It needs more time and expertise as its algorithm-guided questions. In contrast, the Downs & Black checklist facilitates rapid evaluations, reasonable quality and external validity, however, its limited granularity may overlook critical bias domains.
- The findings in the current study suggest that Downs & Black is suitable for preliminary or less rigorous reviews, while ROBINS-I V2 provides a more comprehensive framework for HTA decision-making where methodological robustness is prioritized.

References:
1. ROBINS-I V2 2024 [Available from: https://www.riskofbias.info/welcome/robins-i-v2].
2. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. Journal of epidemiology & community health. 1998;52(6):377-84.
3. Excellence NICE. Appendix H: Appraisal checklists, evidence tables. GRADE and economic profiles. 2022. [Available from: https://www.nice.org.uk/process/pmg20/resources/appendix-h-appraisal-checklists-evidence-tables-grade-and-economic-profiles-pdf-8779777885]
4. SMC. Guidance to submitting companies for completion of New Product Assessment Form (NPAF) 2022 [Available from: https://scottishmedicines.org.uk/media/8515/guidance-on-npaf.docx.]
5. NCPE. Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland 2018 [Available from: https://www.hiqa.ie/sites/default/files/2019-01/Clinical-Effectiveness-Guidelines.pdf.]
6. HAS. Choices in Methods for Economic Evaluation 2020 [Available from: https://www.has-sante.fr/jcms/p_3216041/en/methodological-guidance-2020-choices-in-methods-for-economic-evaluation.]
7. Oikonomidi T, Pavithran T, Kamra S, Pulleddula K, Geddamuri B, Agrawal P, et al. HTA368 Do We Need to Adapt Our Health Technology Assessment (HTA) Systematic Literature Review (SLR) Methods to Comply With the New Joint Clinical Assessment (JCA) SLR Requirements? Value in Health. 2024;27(12):S428.

ISPOR EUROPE 2025, Glasgow, Scotland, UK | 9-12 November 2025