The Role of Surrogacy Analysis in US and Canadian Health Technology Submissions: A Thematic and Temporal Review

HTA40

Tracy Westley, MScPH; Di Wang, MSc; Yu-Chun Chien, MSc; Paul Spin, PhD | EVERSANA®, Value & Evidence, Burlington, ON, Canada

BACKGROUND

- Surrogate endpoints can support accelerated regulatory approval and downstream health technology assessment (HTA) submissions, enabling earlier patient access to novel therapies[1].
- Among agencies, prescriptive validation guidance has been published by IQWiG (2011)[2], NICE Decision Support Unit (2019)[3], and the Member State Coordination Group on HTA (2024)[4].
- Practical guidelines for cost-effectiveness analysis were disseminated by a cross-agency (NICE, CDA-AMC, ICER, ZIN and other countries) working group (2025)[5] and a good practices report on the evaluation of surrogate endpoints was published by an ISPOR Task Force (2026)[6].
- However, empirical characterization focusing specifically on North American HTA bodies and their engagement with surrogate endpoint evidence has been limited.

OBJECTIVE

- To characterize the application and validation of surrogate endpoints across all 2015–2025 **ICER** assessments and **CDA-AMC** reimbursement reviews, including temporal trends and thematic patterns in reviewer critique.

METHODS

- **Study Identification:** All publicly available 2015–2025 CDA-AMC reimbursement reviews (n=689) and ICER assessments (n=98) were identified from their respective websites.
- **Coding Schema:** The extent of surrogate methods validation and reviewer discussion from each CDA-AMC and ICER report were characterized by four dimensions:
  1. **Evidence Classification:** Formal Analysis (conducted *de novo* surrogacy analysis aligned with published validation criteria); Informal Analysis; Assumption (clinical-expert/regulatory precedent); Surrogacy Literature Cited; No Evidence/assumed valid surrogate relationship; Use of a surrogate endpoint not mentioned.
  2. **Collapsed Evidence Category:** *De Novo* Surrogate Analysis (Informal or Formal); Literature and/or Clinical input; No Evidence/assumed valid surrogate relationship; Use of a surrogate endpoint not mentioned.
  3. **Reviewer's Perspective:** No Discussion; Discussion – No opinion/neutral; Discussion – Critical; Discussion – Supportive.
  4. **CDA-AMC Decision Outcome**: Positive Reimbursement Recommendation (any 'reimburse' or 'list' designation, with or without conditions); Do not reimburse.
- Coding schema was conducted using a human-reviewed generative AI classification task implemented in Claude Code with prompting in R.
- Chi-squared tests assessed associations between evidence/reviewer dimensions and CDA-AMC decision outcome (conditional on coding being feasible).
- It was considered not feasible not to apply the Level 1/Level 2/Level 3 validation framework[7] as a coded variable due to unclear reporting within the CDA-AMC and ICER reports.

SURROGATE VALIDATION – STRENGTH OF THE EVIDENCE

L1 **Trial Level:** Evidence is sourced from a meta-analysis of multiple randomized clinical trials. Estimated treatment effects on the surrogate endpoint demonstrate correspondence with estimated treatment effects on the final, patient-centered outcome. This supports the argument that changes in the surrogate endpoint (e.g., increases or decreases) can reliably predict corresponding changes in the final outcome.

L2 **Patient-level:** *Level 2*: Evidence is sourced from interventional, observational (non-interventional), or epidemiological studies as patient-level or cohort data. Estimates of the surrogate endpoint and the final, patient-centered outcome consistently demonstrate an association (e.g., high correlation).

L3 **Biological Plausibility:** *Level 3*: Evidence is limited to findings from pathophysiological studies and/or an established understanding of the disease mechanism. A biologically grounded rationale supports the relationship between the surrogate endpoint and the final, patient-centered outcome. This evidence is qualitative and considered supportive rather than confirmatory of the surrogate endpoint validation.

RESULTS

90 ICER assessments scanned

697 CDA-AMC reviews scanned

42.4% Reports stated use of a surrogate endpoint

<1% Reports included *de novo* surrogacy analysis

Number of Reports Citing Use of a Surrogate Endpoint by HTA body and by Therapeutic Area

Figure 1: By HTA per Year

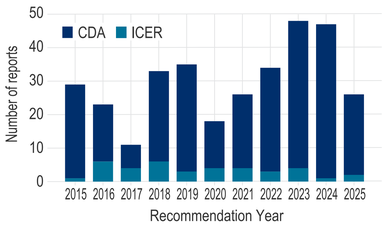
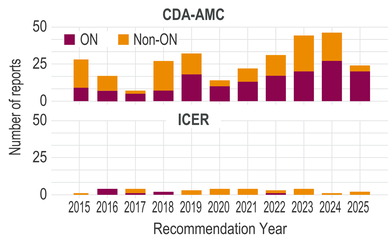


Figure 2: By Therapeutic Area per year



EVIDENCE DOCUMENTED IN PUBLIC-FACING SUMMARY DOCUMENTS

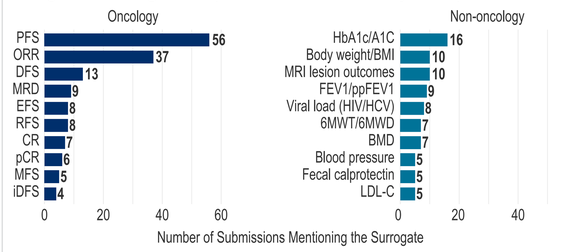
Evidence Category	N (%)
De Novo Surrogate Analysis[a]	7 (0.9%)
Literature Cited/Clinician Input[b]	223 (28.3%)
No Evidence/Assumed[c]	104 (13.2%)
No mention of surrogate endpoint use[d]	453 (57.6%)

a. Refers to reports that include information related to a *de novo* surrogate analysis conducted by the manufacturer and included in a submission. b. Includes narrative arguments, clinical opinion, citation of evidence summarized and/or endorsed by other agencies (e.g., Food & Drug Administration). c. No evidence provided, the surrogacy relationship was assumed to hold. d. Report does not mention use of a surrogate endpoint.

- Across a decade of North American submissions, 334 of 787 (42.4%) mentioned use of a surrogate endpoint.
- Of reports including some evidence to support the surrogacy relationship, the vast majority cited evidence from prior (independent) studies, clinician input, or other examples of acceptance.

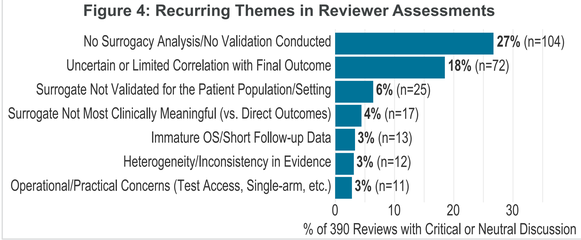
SURROGATE-FINAL OUTCOME PAIRS AND COMMON REVIEWER CRITICISMS

Figure 3: Surrogate Endpoints in Submissions by Therapeutic Area



- Most frequently reported surrogate endpoints were progression-free survival, objective response rate, and disease-free survival, all in oncology, with overall survival serving as the final outcome (**Figure 3**).

Figure 4: Recurring Themes in Reviewer Assessments



- Most frequent reviewer criticisms included: lack of surrogacy analysis; uncertainty about evidence; and lack of validation for the population of interest (**Figure 4**).

CDA-AMC POSITIVE RECOMMENDATIONS BY EVIDENCE CATEGORY AND BY REVIEWER PERSPECTIVE

Figure 5. Proportion with Positive CDA-AMC Recommendation by Evidence Category

Chi-squared p=0.252

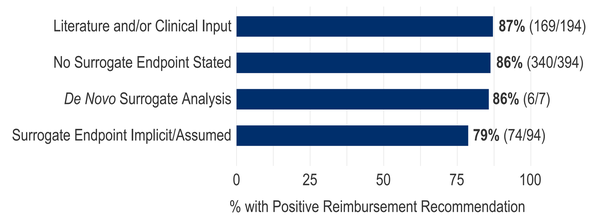
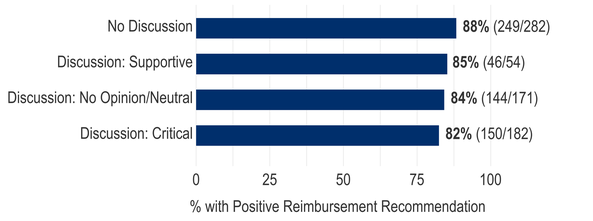


Figure 6. Proportion with Positive CDA-AMC Recommendation by Reviewer Perspective

Chi-squared p=0.333



- Positive CDA-AMC recommendations (e.g., reimburse and/or list with or without recommendations) ranged from 79-87% across all evidence categories (**Figure 5**).
- The lowest positive recommendation rate (82%) was observed when the reviewer perspective was critical of the surrogate evidence (**Figure 6**).
- Neither the classified evidence category nor reviewer perspective were statistically associated with having a positive CDA-AMC reimbursement recommendation (p>0.05).

DISCUSSION

- Surrogate endpoints are often used in ICER and CDA-AMC submissions but applying validated and relevant surrogacy evidence is rare.
- Across 787 North American HTA reports (2015–2025), ~42% included use of a surrogate endpoint and among those reports, ~2% included a sponsor-conducted (*de novo*) surrogacy analysis.
- Too few instances of surrogate validation exercises were identified to statistically test temporal trends in the uptake of the various published guidelines.
- It was not always clear from the reports the exact methods of surrogacy analysis, either performed *de novo* or cited as supportive publications. Therefore, the formal Levels of Evidence (L1/L2/L3) criteria could not be applied.
- Neither evidence category nor reviewer perspective of the surrogacy evidence was statistically associated with the reimbursement outcome.
- When CDA-AMC reviewer engagement with the use of a surrogate endpoint was clearly directional, criticism outnumbered support roughly 3.3 to 1.

CONCLUSIONS

- North American reviews and assessments mentioning the use of surrogate endpoints provide some but limited reporting on the type of evidence provided, how it informs the models (e.g., clinical efficacy and/or cost-effectiveness), and the extent that the surrogacy methods (or lack of) influence reviewer recommendations and decisions.
- While more explicit and practical guidance on the validation and use of surrogate methods in cost-effectiveness has been recently published, it would benefit future sponsors and submission teams to be able to learn from precedent reports.
- More transparency in reporting, with increased attention to the characteristics of the surrogacy evidence provided and any influence on decision-makers, is recommended for future submissions.

ABBREVIATIONS: AI = artificial intelligence; BMD = bone mineral density; BMI = body mass index; CDA-AMC = Canada's Drug Agency-L'Agence des médicaments du Canada; CR = complete response; DFS = disease-free survival; EFS = event-free survival; FEV1 = forced expiratory volume at one second; HbA1c = hemoglobin A1c; HCV = hepatitis C virus; HIV = human immunodeficiency virus; HTA = Health Technology Assessment; ICER = Institute for Clinical and Economic Review; iDFS = invasive DFS; IQWiG = Institute for Quality and Efficiency in Health Care; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; LDL-C = low-density lipoprotein cholesterol; MFS = metastasis-free survival; 6MWT = 6-minute walk test; MRD = minimal residual disease; MRI = magnetic resonance imaging; NICE = National Institute for Health and Care Excellence; Non-ON = non-ON; ON = oncology; ORR = overall response rate; OS = overall survival; pCR = pathological CR; PFS = progression-free survival; pp = percent predicted; RFS = relapse-free survival; US = United States; ZIN = Zorginstituut Nederland.

REFERENCES: **1.** Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics (2018). *FDA*. **2.** Validity of surrogate endpoints in oncology: Executive summary of rapid report A10-05, Version 1.1. (2011) *IQWiG*. **3.** Bujkiewicz et al. NICE DSU TSD20: Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints. 2019. **4.** Guidance on outcomes for joint clinical assessments (2024). *MS CG on HTA*. **5.** Surrogate endpoints in cost-effectiveness analysis for use in health technology assessment (2025). *ICER, CDA-AMC, NICE, ZIN*. **6.** Bujkiewicz et al. (2026). Methods for Evaluation of Surrogate Endpoints for HTA Decision Making: A Good Practices Report of an ISPOR Task Force. Value in Health. **7**. Ciani et al. (2016). Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. V*alue in Health*. 20(3): 487-495.

ACKNOWLEDGMENTS: The authors thank Kelsey Guenther (data extraction and editorial review) and Elizabeth Willis (editorial and poster support) for their contributions.

CONTACT: Tracy.Westley@eversana.com

Presented at ISPOR US 2026
Philadelphia, PA, USA
May 17 to 20, 2026

EVERSANA®