

How Are ICER's Evidence Ratings Determined? A Systematic Review of ICER's Evidence Ratings in Evidence Reports for New Drugs in 2020 and 2021

Naoko A. Ronquest, PhD; Micah Clinkscales, MBA; Kyle Paret, PhD
RTI Health Solutions, Research Triangle Park, United States

BACKGROUND

- The United States Institute for Clinical and Economic Review's (ICER's) value assessment framework is designed to align with methods used by major global health technology assessment agencies.
- However, ICER's method for rating evidence for each intervention's comparative clinical effectiveness is unique.¹

OBJECTIVE

- To understand (1) how evidence ratings were assigned to interventions recently reviewed by ICER and (2) the types factors that may influence ICER's ratings decision.

METHODS

Data Source, Review Structure, and Study Period

- Based on a systematic review of all evidence reports for pharmacotherapies published in 2020 and 2021, we summarized characteristics of interventions for each clinical rating and the frequency of rating revisions between draft and final reports.

Outcomes Extraction Strategy

- For each assessment, the following characteristics were extracted:
 - Interventions reviewed
 - Comparator selected for each clinical rating
 - Subpopulation selected for each clinical rating
 - Whether rating was given in an "update" review
 - Therapeutic area categories (rare disease, chronic disease, acute disease)

- For each pharmacotherapy, the following characteristics were extracted:

- ICER's draft and final clinical ratings
- Whether the ratings changed between draft and final assessments
- Whether the comparator was an active drug or a best supportive care

Critical Assessment

- For each assessment that involved a change in rating, we summarized potential reasons for the influence of stakeholder comments on the rating.

RESULTS

Overview of Reviewed Assessments in 2020 and 2021: Therapeutic Areas and Treatments Reviewed

- In total, 45 interventions were reviewed across 17 assessments published in 2020 and 2021 (2020, n = 7; 2021, n = 10).
- Types of assessments:
 - Therapeutic areas: Chronic diseases: 82% (14/17); acute diseases: 18% (3/17); rare diseases: 47% (8/17)
 - Assessment update: 24% (4/17)
- Most ICER reports presented multiple clinical ratings for each assessed intervention, depending on the number of subpopulations and comparators evaluated (Figure 1).
 - 8 assessments provided separate ratings for each intervention (6/17 for different subpopulations, 8/17 for different comparators).
 - 68 total ratings for combinations of interventions, populations, and comparators.

Highlights of ICER's Clinical Ratings

- Although many assessments resulted in either promising but inconclusive (P/I: 19% [13/68]) or insufficient (I: 19% [13/68]) ratings, more than one-third of cases were B+ or better (B+: 24% [16/68]; A: 13% [9/68]).

Potential Determinants of Ratings

- Figure 3 summarizes all ratings given to assessed interventions in all selected comparator/population combinations in 2021 and 2020.
- Ratings of B+ were usually associated with a sizeable improvement in clinical outcomes without longer-term safety evidence (or less impressive efficacy with longer-term safety evidence).
 - Multiple myeloma assessment in 2021 (CAR-Ts: B+): CAR-Ts' overall response, 60%-75% vs. 30% (standard of care)
 - Lupus nephritis assessment in 2021:
 - Voclosporin (B+): Nearly doubled complete response vs. standard of care (background therapy only) (42.3% vs. 23.3%) at 12 months
 - Belimumab (B+): Higher complete response rate vs. standard of care (32.5% vs. 25.5%) but not as large a difference as voclosporin vs. standard of care; longer-term safety (2-year follow-up) results were published before ICER's review
- In A ratings, there typically was a sizable improvement in clinical outcomes and longer-term safety follow-up.
 - Example: A review of treatments for hereditary angioedema (HAE) in 2021
 - The percentage reduction in total HAE attacks was estimated to be excellent for all 3 interventions (50%-90%, statistically significant) that received the A rating.
 - 1 treatment received the P/I rating because the treatment was a new class of "biologic" treatment without long-term safety follow-up.
- When only indirect efficacy evidence was available, the ratings were C+ or below (e.g., ulcerative colitis, 2020).

Results of Critical Assessment of Stakeholder Comments

- There were only 4 instances (for 5 interventions) where evidence ratings changed between draft and final reports (Table 2).
 - 4 rating revisions resulted in improved ratings, and 1 revision resulted in a lower clinical rating.
- 1 rating was revised spontaneously (unrelated to stakeholders' public inputs) when ICER identified new evidence after publishing a draft report (high cholesterol).
- In another revision (hemophilia A), ICER's incorporation of stakeholder comments resulted in a lower rating of 1 of the treatments reviewed.

Figure 1. Characteristics of Reviewed ICER Assessments in 2020 and 2021

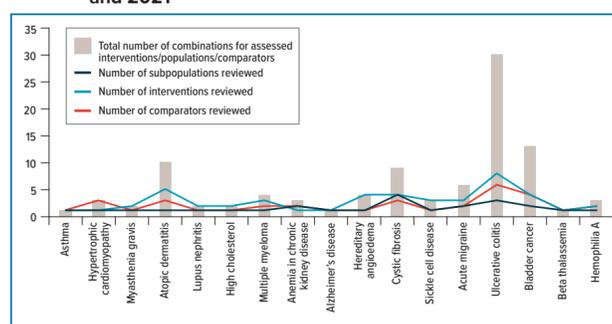


Figure 2. Summary of Final ICER Ratings in 2020 and 2021

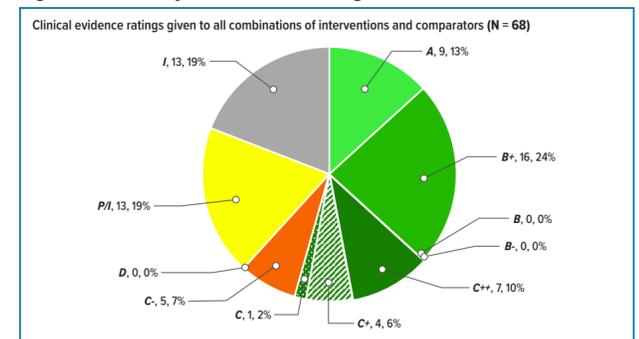


Figure 3a. Clinical Evidence Ratings of Interventions Evaluated in 2021

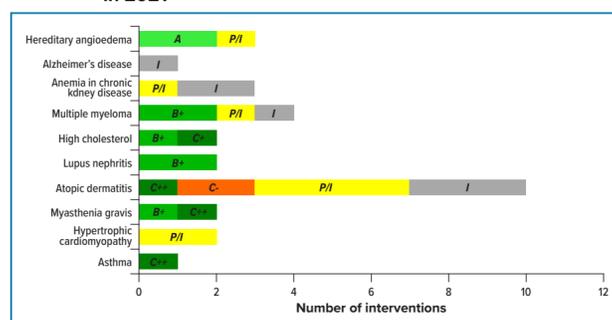


Figure 3b. Clinical Evidence Ratings of Interventions Evaluated in 2020

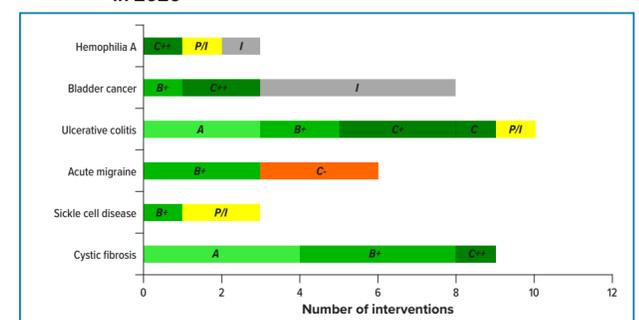


Table 1. Summary of Rationale in Evaluations that Changed Their Clinical Ratings Between Draft and Final Reports

Therapy area	Intervention with rating change	Rating change	Reasons for change	Stakeholder influence?
High cholesterol (2021)	Inclisiran vs. placebo	C++ → B+	Availability of longer-term safety data (nearly 2 years of data showing no significant AEs)	Unlikely <ul style="list-style-type: none"> No comment pertaining to safety data that led to inclisiran's rating improvement
Acute migraine (2020)	Lasmiditan vs. sumatriptan and vs. eletriptan	D → C-	New NMA data that showed all interventions to be more efficacious vs. placebo	Likely <ul style="list-style-type: none"> Significant improvement in NMA outcomes (active vs. placebo) for all interventions led to improved clinical ratings for all treatments
Bladder cancer (2020)	Nadofaragene firadenovec vs. BSC and oportuzumab monatox vs. BSC	C+ → C++	To reflect the possibility of a substantial benefit	Yes, a stakeholder requested a change in rating from C+ to B+ <ul style="list-style-type: none"> Clinical evidence demonstrating that a new treatment results in clinically significant improvement on at least 1 endpoint was submitted
Hemophilia A (2020)	Emicizumab vs. factor VIII prophylaxis	B+ → C++	Higher real-world dose was used in the final report, which led evaluated interventions to be associated with additional efficacy, although the magnitude of the additional efficacy was uncertain.	Possible <ul style="list-style-type: none"> Various stakeholders including patient groups recommended ICER to incorporate real-world dosing of treatments

AE = adverse event; BSC = best supportive care; NMA = network meta-analysis.

CONCLUSIONS

- Although a considerable portion of evidence ratings in recent ICER reports were B+ or better, stakeholder inputs rarely made a difference in ratings.
- Future research is warranted to better characterize and quantify the health benefit (magnitude and likelihood) needed to achieve each rating.

REFERENCES

- ICER. 2017. <https://icer.org/wp-content/uploads/2020/10/Rating-Matrix-User-Guide-UPDATED-06.30.17.pdf>.

FUNDING

No funding to disclose

CONTACT INFORMATION

Naoko Ronquest
Executive Director, US HTA Strategy and Early Modeling
RTI Health Solutions
Email: nronquest@rti.org