# THE QUALITY OF HEALTH ECONOMIC STUDIES FOR ECONOMIC EVALUATIONS OF CLINICAL PHARMACY SERVICES: A QUALITATIVE ANALYSIS

Kim M PharmD[1], De Los Santos B PharmD MS[1], Alsharif A PharmD[2], Perez A PharmD MS[2], Touchette D PharmD MA[1]

[1]UIC College of Pharmacy, Department of Pharmacy Systems, Outcomes, and Policy
[2] Nova Southeastern University Florida Barry and Judy Silverman College of Pharmacy

PHARMACY SYSTEMS OUTCOMES AND POLICY COLLEGE OF PHARMACY — UIC

NSU Florida — Barry and Judy Silverman College of Pharmacy — NOVA SOUTHEASTERN UNIVERSITY

## Supplement

## METHODS

- A systematic review was conducted to assess economic evaluations of clinical practice studies using the Quality of Health Economic Studies (QHES) instrument. Following this review, a survey comprising four open-ended questions was developed. The survey aimed to evaluate several key aspects of the QHES: the clarity and interpretation of its questions, its effectiveness in generating consistent scores, its overall capability to capture essential constructs of study quality, and its overall capability to measure the quality of economic evaluations. Thematic analysis was employed to analyze the responses to each open-ended question. This analysis sought to identify and categorize key themes that emerged from the participants' feedback. The identified themes were grouped into four main categories: (1) Subjectivity/Ambiguity of Questions, (2) User Experience, (3) Scoring Limitations, and (4) Other Improvements. To assess the reliability of the survey responses, interrater reliability was measured using Cohen's Kappa for categorical agreement between raters. Additionally, the consistency of QHES scores across different raters was evaluated using the Intraclass Correlation Coefficient (ICC). These statistical methods ensured the robustness and reliability of the insights gathered from the survey.

## METHODS

### Survey Questions

Q1. Can you discuss any challenges you encountered when using the QHES to evaluate health economic studies? Specifically, were there any questions within the QHES that you found particularly difficult to answer or for which it was unclear what the question was asking? If so, please explain.

Q2. Reflecting on your use of the QHES, what are your thoughts on its ability to generate consistent overall scores for health economic studies? Do you have any concerns about the results you obtained?

Q3. A study's quality is typically assessed by the following constructs:
Can the study be believed? (assessment of internal bias)
Are the results relevant? (assessment of external bias)
Is there the possibility of variation in the study's findings? (assessment of adherence to protocol)
Are the study methods and results clearly communicated? (assessment of transparency)

In your opinion, how effectively does the QHES instrument capture these essential constructs of a study's overall quality and validity with regards to the health economic assessment? If there were missing constructs that you thought were important or constructs that were not critical to a study's design, describe what these omissions or extraneous information were.

Q4. In your opinion, how well do you think the QHES measures the quality of economic evaluations?
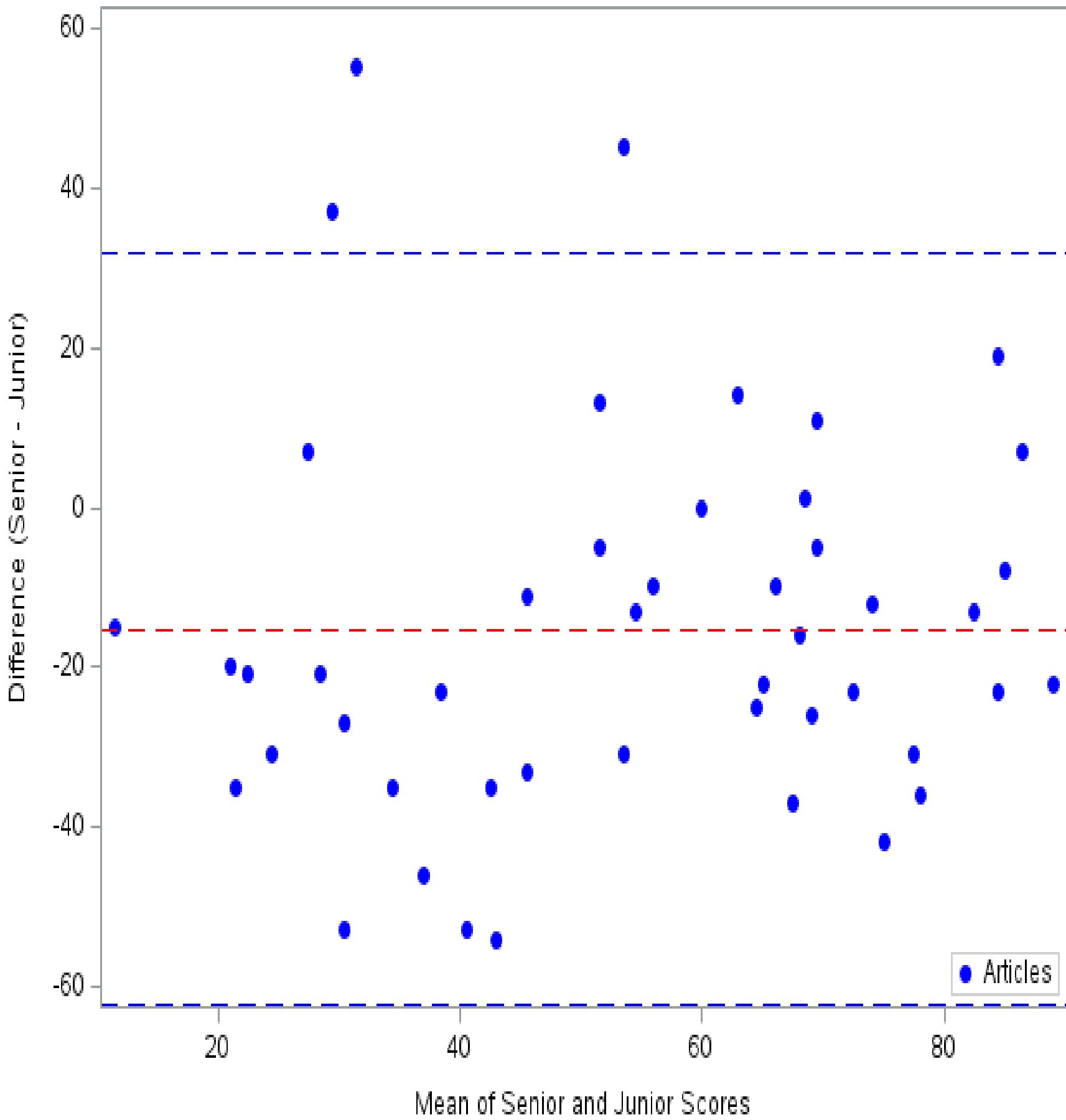
## RESULTS



Figure 2: Cohen's Kappa per QHES Question
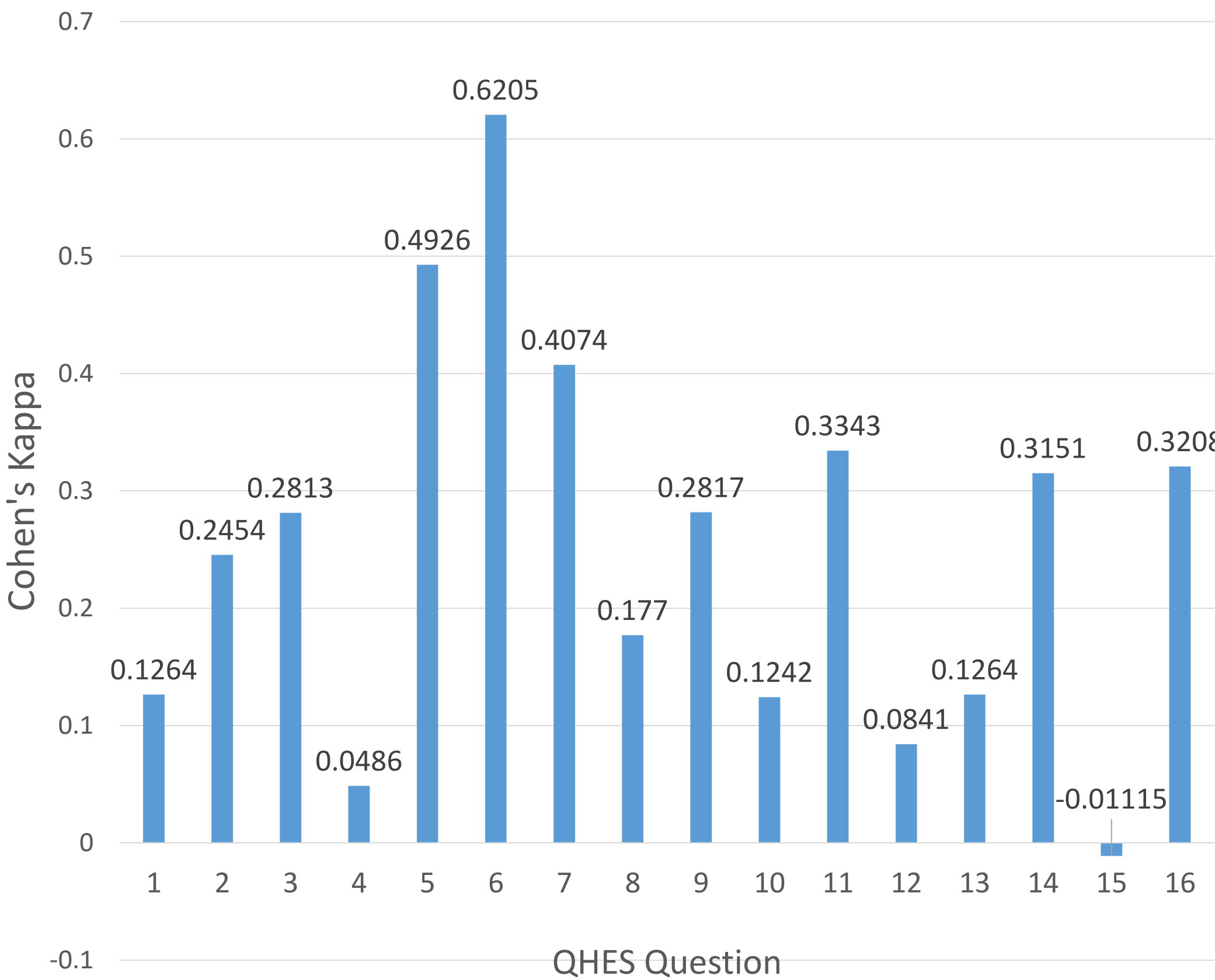
### Kappa Summary
- **None (0–0.20):** 7 (QHES 1, 4, 8, 10, 12, 13, 15)
- **Minimal (0.21–0.39):** 7 (QHES 2, 3, 9, 11, 14, 16)
- **Weak (0.40–0.59):** 2 (QHES 5, 7)
- **Moderate (0.60–0.79):** 1 (QHES 6)
- **Strong (0.80–0.90):** None



Figure 3: Bland-Altman Plot of Senior vs Junior QHES Scores

## RESPONSES

### Subjectivity/Ambiguity of Questions

- ...wording of the questions was hard to understand. For example, Q3 "were variable estimates used in...", what variable estimates does this refer to? Cost estimates from the pharmacist salary wage or time spent providing care? Because while a clinical trial may give the best estimate for time spent providing care I would say the best salary estimate would come from a publicly available source like US Bureau labor statistic data.

- Q4 "If estimates came from a subgroup..." what estimates does this refer to? Result estimates coming from a subgroup analysis are definitely not as valid when they are not specified a priori but based on the wording of the questions

- I feel uncertain that the QHES tool can generate consistent overall scores, especially between multiple reviewers. I try to refer back to our QHES answering guidance when it's unclear on how I should answer but this answering guidance was developed internally and reviewers outside of our team would not necessarily have answered the same way. The fact that we had to develop internal answering guidance in fact speaks to the lack of consistency in using the tool and our team's method of overcoming that.

- ...if these details are not explicitly provided in the study, it may lead to lower scores due to perceived gaps in reporting, rather than actual deficiencies in study quality

### User Experience

- One of the challenges I encountered was with questions related to the methodological rigor of the studies. For example, the question about whether the study design and the analysis methods used were appropriate for the research question can be challenging. It requires a deep understanding of various study designs and their appropriateness for different types of economic evaluations, which can be difficult without extensive background knowledge.

- "Questions can be interpreted in very different was based on your level of experience, and even among those with high skill level."

- "For many of the questions (eg: Q1, Q5, Q9, Q10, Q11), the reviewer's skill in assessing the overall quality of each aspect of the study might lead to different responses."

- There is likely a high probability of generating different scores based on the reviewer's experience with economic analyses

### Scoring Limitations

- In several cases, there are multiple questions in one and a yes/no answer option is not appropriate

- Most of the questions were not applicable to the studies I included. Because our systematic review included partial economic evaluations many of the questions (6,12,13) I could not answer as the QHES tool was made to assess full economic evaluations  (and feels specifically geared towards modeling economic evaluations).

- Scoring Bias: The weighted nature of the QHES scoring system means that certain aspects of a study are deemed more important than others. While generally reflective of the critical elements of health economic evaluations, this can lead to skewed outcomes if a study scores poorly on heavily weighted questions despite performing well on others. In addition, the yes/no interpretation to multilevel questions may not accurately reflect the overall study's quality (or lack thereof) as there may be some "gray" area in between yes/no.

- All questions in QHES are generally crucial for assessing health economic evaluations. However, the weight (points) given to some questions might not proportionately reflect their importance in every study scenario, leading to potential overemphasis or underemphasis of certain aspects based on the scoring system.

- Scoring of certain questions (Q4 in particular) may be too specific and not applicable for most studies.  Having only two response levels for each question may make assessment of quality even less reliable.

- Potential Overemphasis on Certain Aspects: The weighted nature of the QHES might lead to overemphasis on certain aspects of a study while potentially underemphasizing others . For instance, if a study performs exceptionally well in heavily weighted areas but poorly in others, its overall quality might be misrepresented.

### Other Improvements

- Furthermore, the analysis does not address the relative importance of potential biases on the study results.

- Detailed Guidelines: Develop more detailed guidelines for each question, including examples of what constitutes a high or low score.

- The QHES could be enhanced by incorporating elements such as patient-centered outcomes, equity and ethical considerations, and long-term impacts and sustainability of interventions.

- Improvements could be made so that assessment of study methods and accuracy of results are better measured

- ...it could provide more explicit guidelines and criteria for evaluating the external validity or generalizability of the results, which is crucial for understanding how the findings apply in real-world settings.