

What Impacts the Quality of Comparative Effectiveness Research: A Classification and Regression Tree Analysis Using the GRACE Checklist

Allison Bryant¹, Aaron B. Mendelsohn², Shilpa Viswanathan², Nancy A. Dreyer¹
¹Quintiles, Cambridge, MA, USA; ²Quintiles, Rockville, MD, USA

Objectives

The Good ReseArch for Comparative Effectiveness (GRACE) checklist is a tool for evaluating the quality of comparative effectiveness research (CER) studies. The checklist consists of 11 questions on data and methods and was developed through literature review, expert consultation, and testing by 113 raters across five continents. The purpose of the present research was to determine which checklist questions are most predictive in identifying quality CER.

Methods

- Twenty-two volunteers recruited from academia, industry, and government applied the GRACE checklist to 28 CER articles, for a total of 56 assessments.
- Classification and Regression Tree (CART) analysis, a binary, recursive, partitioning methodology, was used to identify the checklist questions that were most predictive of quality CER articles.
- Quality was defined using four external measures of quality:
 - Journal impact factor >2.5
 - Article citations per year >2
 - Expert assessment of overall quality classified as “sufficient”
 - Composite outcome of all three quality measures
- CART default settings were altered to vary the penalty for misclassifying sufficient versus insufficient quality articles.

Results

- The use of a composite outcome in the CART analysis yielded on average a higher sensitivity and specificity than any of the outcomes individually.
- Altering the penalty for misclassifying sufficient versus insufficient quality articles had minimal impact on the sensitivity and specificity.
- The use of sensitivity analysis was the strongest predictor of quality.

Table 1. Performance of CART Algorithms Using GRACE Checklist Items in Predicting Quality Outcomes

	Sensitivity ¹	Specificity ²
Expert Assessment	78.1%	58.3%
Article Citations	50.0%	63.3%
Journal Impact Factor	66.7%	85.3%
Composite Outcome	71.4%	80.9%

¹Sensitivity indicates the proportion of articles with sufficient quality outcomes that were deemed sufficient by the volunteer raters. ²Specificity indicates the proportion of articles with insufficient quality outcomes that were deemed insufficient by the volunteer raters.

Figure 1. Expert Quality Assessment*

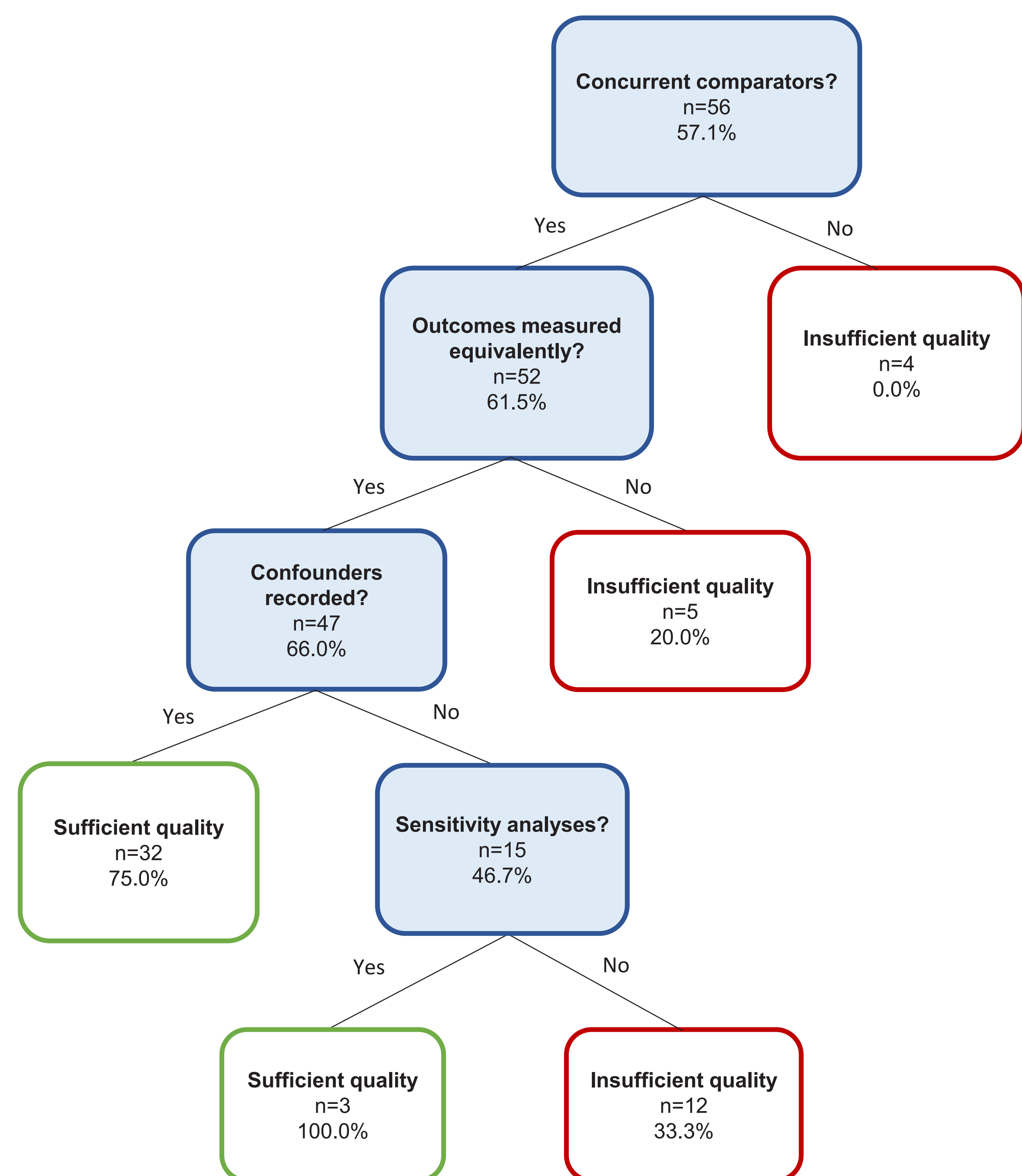


Figure 2. Number of Article Citations per Year*

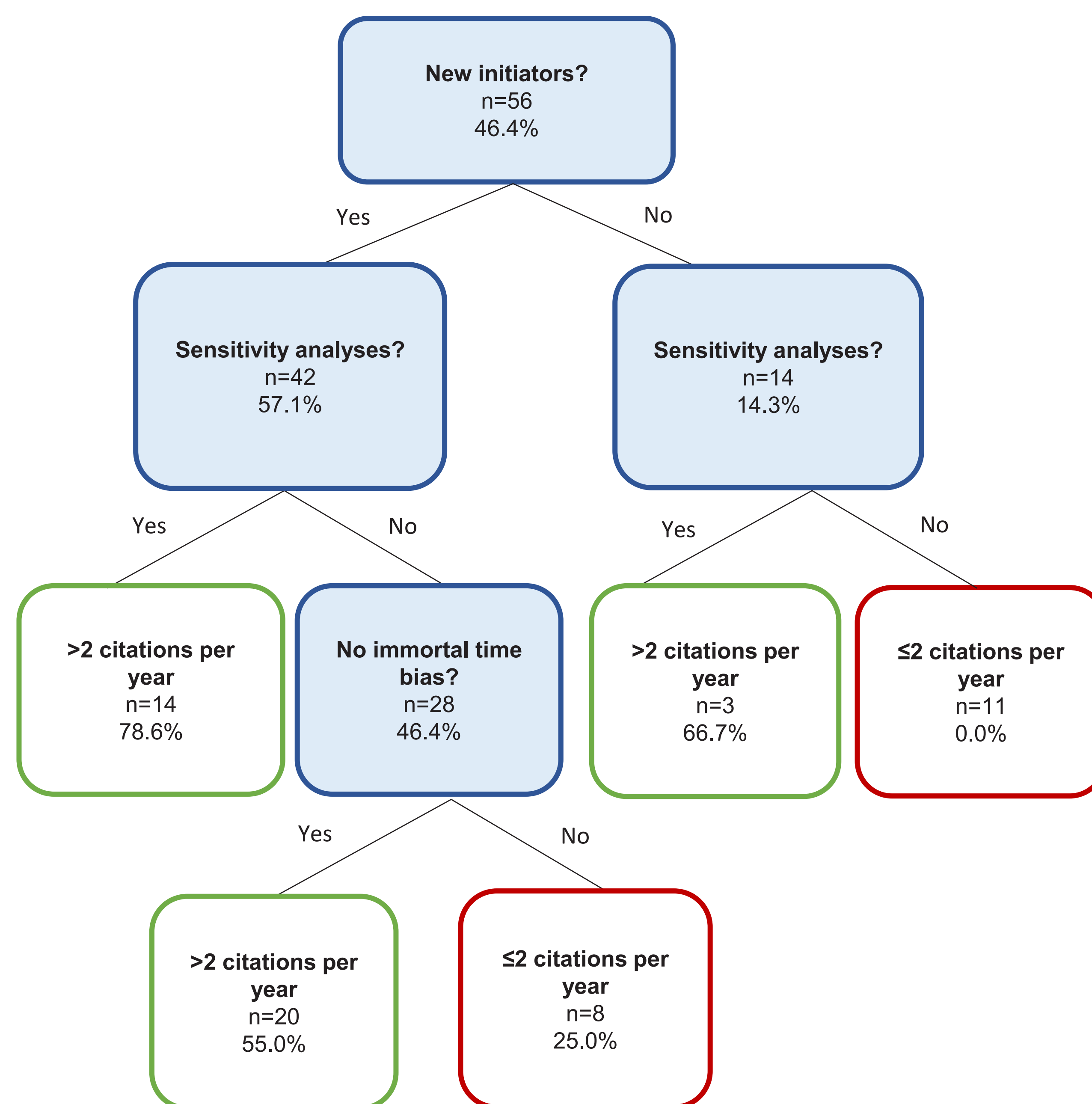


Figure 3. Journal Impact Factor*

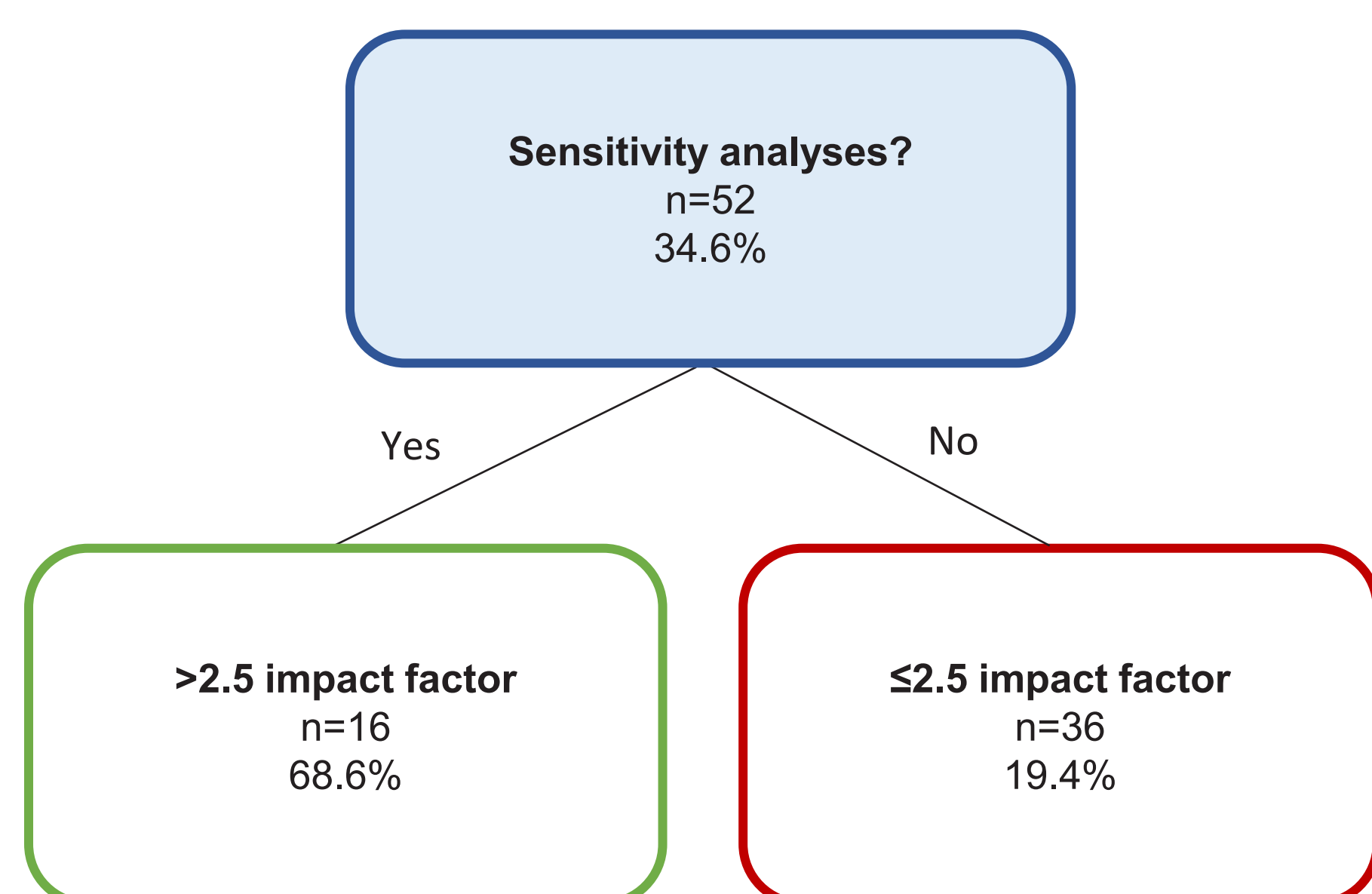
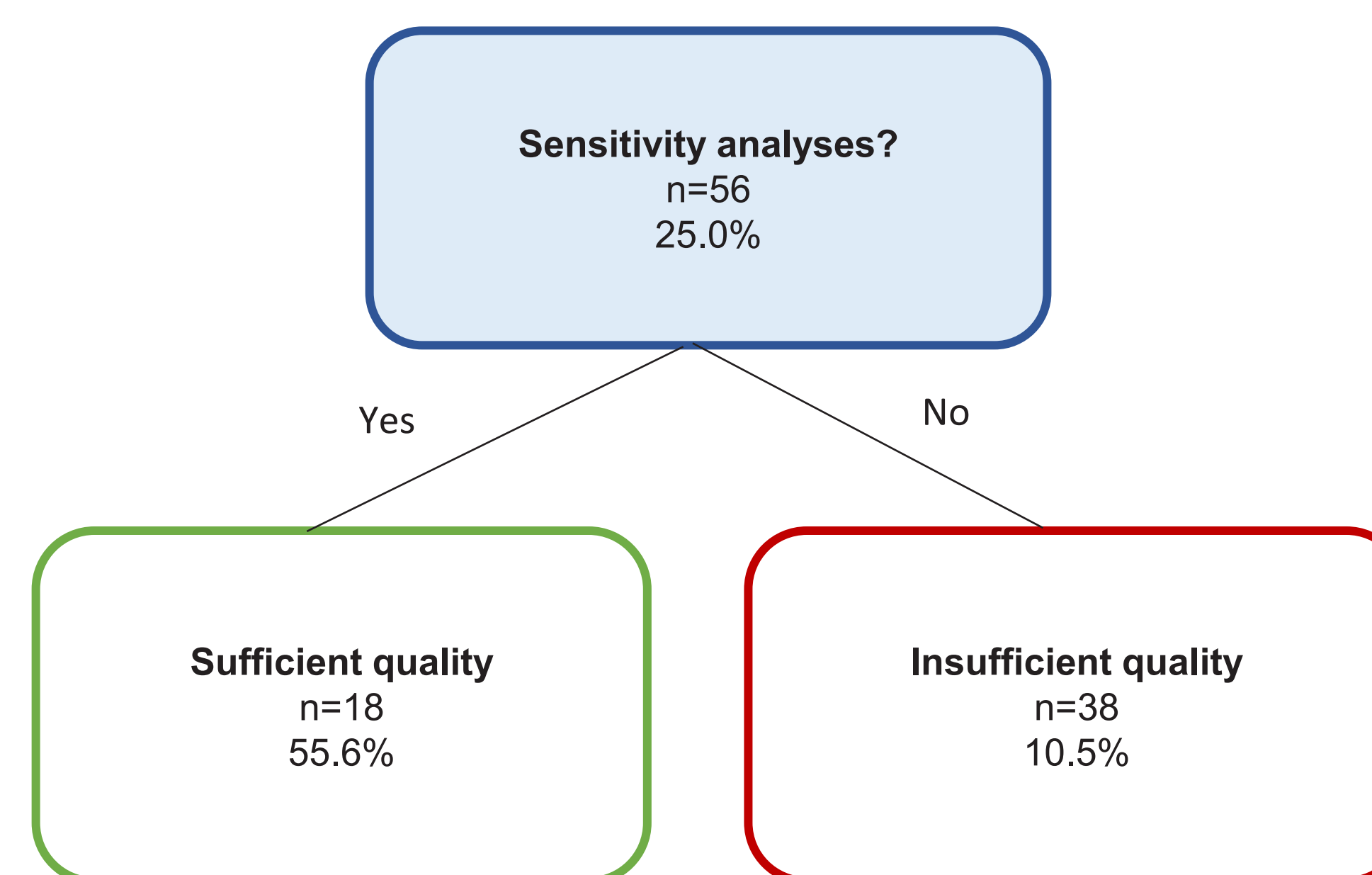


Figure 4. Composite Outcome*



*n is the total cases in each node; % is the proportion of cases classified as having a sufficient expert quality rating, >2 article citations per year, >2.5 journal impact factor, or sufficient composite outcome, respectively

Conclusion

When the GRACE checklist for assessment of the quality of observational CER studies was applied to 28 articles and compared with four external measures of article quality using CART analysis, the strongest predictors of quality included: use of concurrent comparators; limiting the population to new initiators of the study drug; equivalent measurement of outcomes across treatment groups; reasonable collection of data on confounders; accounting for immortal time bias; and use of sensitivity analyses to test how much effect estimates depended on various assumptions. Use of sensitivity analyses in particular was found to be a strong predictor of each of the four measures of quality. The composite outcome had overall higher performance compared to the individual quality measures alone.