Serious reporting deficiencies exist in minimal important difference studies: current state and suggestions for improvement

Carrasco-Labra A, Devji T, Qasim A, et al. (2022) Serious reporting deficiencies exist in minimal important difference studies: current state and suggestions for improvement. *J Clin Epidemiol*; 150:25-32.10.1016/j.jclinepi.2022.06.010.

Lertkong Nitiwarangkul, MD.

Objectives

- To evaluate reporting of minimal important difference (MID) estimates using anchor-based methods for patient-reported outcome measures (PROMs)

- Association with reporting deficiencies on their credibility.

Methods

- Systematic survey of primary studies empirically estimating MIDs
  - Medline, EMBASE, PsycINFO, and the Patient-Reported Outcome and Quality of Life Instruments Database
  - Until October 2018
- Evaluated study reporting
  - Participants' demographics, intervention(s), characteristics of PROMs and anchors, and MID estimation method(s).
- Assessed the impact of reporting issues on credibility of MID estimates

Briefly summarize

- In 585 studies / 5,324 MID estimates / 526 distinct PROMs
- Authors frequently failed to adequately report
  - Characteristics of PROMs and MIDs, including
  - Minimum and maximum,
  - Measure of variability
- Most serious reporting issues impacting credibility included
  - Infrequent reporting of the correlation between the anchor and PROM (66%)
  - Inadequate details to judge precision of MID point estimate (13%)
  - Insufficient information about the threshold used to ascertain MIDs (16%)

Background

- High quality reporting is essential to inform users
- Limitations in reporting threaten users' ability to effectively evaluate research findings
- Nearly 50% of research reports suffered from serious issues that made them virtually unusable
- Anchor-based minimal important difference (MID)
  - Point of important change in health status
  - Small but important

Background

- The PRO extension of the Consolidated Standards of Reporting Trials Statement (CONSORT PRO), published in 2013
- Proposes a number of items relevant to make RCT reports including PROs more informative
- Include discussion of a minimal important change or a responder definition
- Lack of reporting standards
  - Great variability in reporting
  - Users of MIDs may not be able to effectively evaluate their credibility

Rationale of conduct this study

- To evaluate the reporting in studies empirically calculating MID estimates using anchor-based methods
- And inform future reporting standards

- **This study undertook a systematic evaluation of reporting in MID estimation studies**

Methods

- We conduct systematic survey include
  - Inventory of available **anchor-based MID estimates**
  - **Instrument** to assess the credibility of these estimates

Selection criteria

- Inclusion criteria
- Primary studies → Calculated anchor based MID for PROMs
- Enroll Adolescents and adults
- All types of anchor instrument used
  - Self-reported
  - Proxy reported
  - Laboratory data
  - Performance-based measure
- Exclusion criteria
  - Systematic review studies
  - Targeted a moderate or large important difference
  - MIDs estimated using a combined anchor and distribution-based approach

Literature search

- Medline, EMBASE, and PsycINFO
  - Until October 2018 (From 1989, year that MID first develop)

- Accessed the Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID)

Study selection

- Use two reviewers independently extract data
- Supervisor will take action when disagreement occur

Data collection and evaluation of reporting quality

- Study characteristics and participants demographics
  - Countries, Number of participant, Eligibility, Record if unclear reported
- Reporting of interventions applied when estimating MIDs
  - Extracted description and recorded if unclear
- PROM
  - Name, Number of domain
  - Minimum and Maximum value (higher or lower values represented a better health state)

Data collection and evaluation of reporting quality

- Anchor
- Description; Constructs, Nature
- Response options
  - Transition rating anchors with response options such as
  - Great deal worse, somewhat worse, a little worse, no change
  - Little better, somewhat better, a great deal better
- Specific threshold
  - Patients who reported feeling ‘‘a little better and somewhat better’’
  - One category change in an 11-point VAS of pain

Data collection and evaluation of reporting quality

- MID determination
- Evaluated whether the authors reported
  - Endpoints considered for MID estimation
  - Length of follow-up between the administration of the PROM and the anchor
  - Analytical approach (e.g., mean change, mean difference, ROC, regression analysis)
  - MID reflects improvement, deterioration, or both
  - Measure of variability for the MID point estimate (e.g., confidence interval, interquartile range, standard deviation, range)

Data collection and evaluation of reporting quality

- Credibility assessment
- Using a new instrument (Table)
- Five core items
  - Ratings of ‘‘definitely no’’ and ‘‘impossible to tell’’ to reflect the lowest levels of credibility
  - Minimize the chance of error, a third reviewer (ACL, TD) served as quality control

Credibility assessment

- Credibility instrument → Instrument has five criteria

Core criteria		
Is the patient or necessary proxy responding directly to both the PROM and the anchor?	Yes	No/impossible to tell
Is the anchor easily understandable and relevant for patients or necessary proxy?	Definitely yes/to a great extent	Definitely no/not so much/impossible to tell
Has the anchor shown good correlation with the PROM?		
Is the MID precise?		
Does the threshold or difference between groups on the anchor used to estimate the MID reflect a small but important difference?		

The result: definitely yes → No concern about credibility

Results

Search results

- 2,161 in full text → 585 studies proved eligible → Reported 5,324 MID estimates for 526 PROMs

Reporting items

- Participants' demographics
  - Single concern → 16% of studies failed to report the country
- Interventions applied when estimating MIDs
  - 13% failed to described the intervention applied
- PROM
  - 22% failed to report the minimum and maximum values of the PROM scale
  - 18% of authors failed to report the meaning of higher values on the PROM scale

Reporting items

- Anchor
  - 6% failed to reported description, response option of anchor and threshold used to define the MID

- MID determination
  - Almost 2/3 of studies didn’t report measure of variability
  - 1/4 of studies failed to report N of participants
  - 64% did not report the correlation between the PROM and anchor

Reporting item	Not reported *n* (%)
Participants' demographics	
Country where the study was conducted	93 (16)
Disease or condition of the participants	0 (0)
Number of participants at baseline	5 (1)
Participants' age measure of central tendency	44 (8)
Participants' age measure of dispersion	54 (9)
Male/female ratio	35 (6)
Intervention(s) applied when estimating MIDs	
Description of the intervention applied	76 (13)
PROM instrument	
Construct measured (total instrument or domain)	15 (3)
Lower and upper values of the PROM scale	124 (22)
Meaning of the extreme values of the PROM	100 (18)
Anchor instrument	
Construct measured	29 (5)
Nature of the anchor (e.g., transition rating, change in disease related outcome)	0 (0)
Source of anchor information	5 (1)
Description of the range of options/values	36 (6)
Description of the threshold used to define the MID	29 (5)
MID estimation	
Length of follow up	45 (8)
Correlation between anchor and PROM	373 (64)
Analytical method	9 (2)
Measure of variability of MID (e.g., CI, IQR, SD, SE, range)	383 (65)
Number of participants included in MID calculation	147 (25)

Abbreviations: MID, minimal important difference; PROM, patient reported outcome; CI, confidence interval; IQR, interquartile range; SD, standard deviation; SE, standard error.

Credibility assessment and association with reporting

Core criteria		
Is the patient or necessary proxy responding directly to both the PROM and the anchor?	Yes	No/impossible to tell
Is the anchor easily understandable and relevant for patients or necessary proxy?		
Has the anchor shown good correlation with the PROM?		
Is the MID precise?	Definitely yes/to a great extent	Definitely no/not so much/impossible to tell
Does the threshold or difference between groups on the anchor used to estimate the MID reflect a small but important difference?		

Largest proportions of MIDs evaluated

Credibility assessment and association with reporting

- Most serious failure of reporting
  - 3,514 of 5,324 (66%) of MID estimates did not include the **correlation** between the anchor and PROM of interest
- Second most serious reporting
  - **Precision** of the MID estimate
  - 13% were judged as ‘‘impossible to tell’’
- Third criterion of concern
  - **Threshold** used to calculate the MID reflects a small but important difference
  - 16% of the MID estimates presenting serious reporting issues

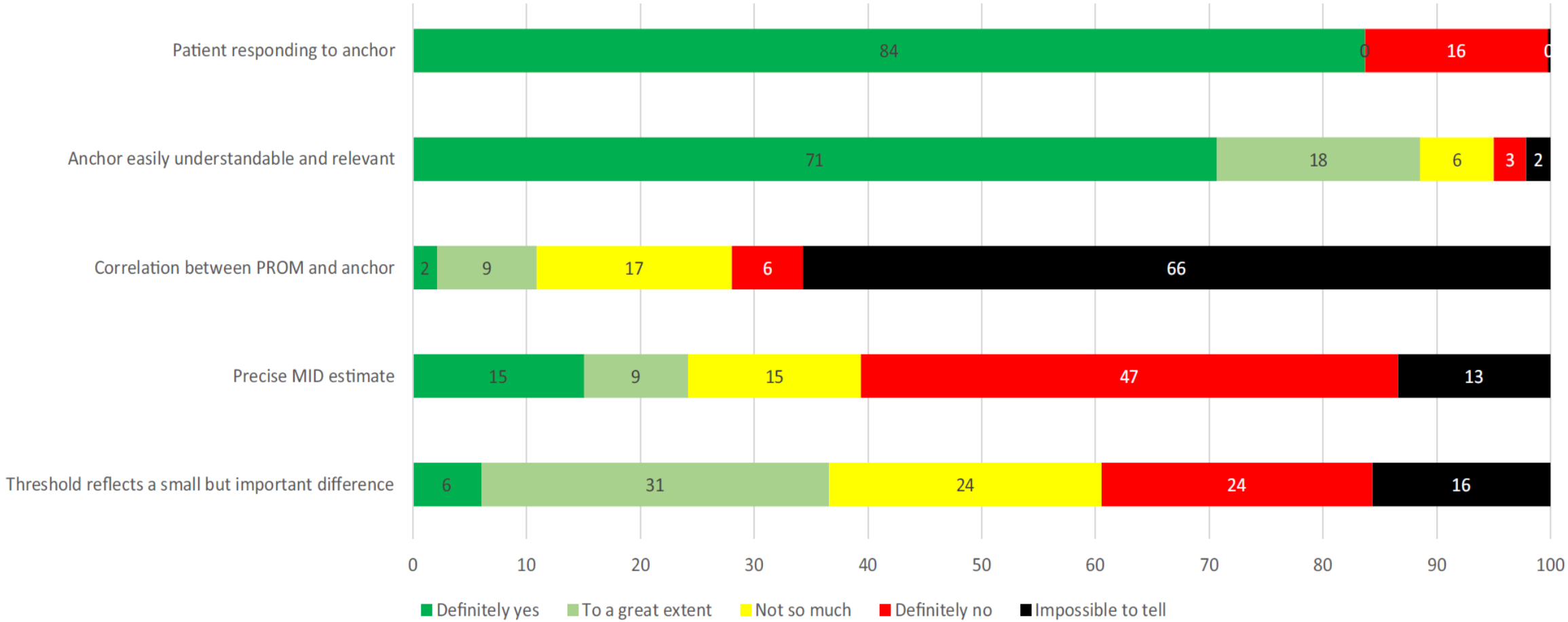

Fig. 1. Proportion (%) of minimal important difference estimates and evaluation of core credibility criteria (n = 5,324 MIDs). MID, minimal important difference; PROM, patient reported outcome measure.(For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Discussion

Main findings

- The most serious issues impacting credibility assessments
  - Infrequent reporting of the correlation between the anchor and PROM
  - Inadequate details to judge precision of the MID point estimate
  - Insufficient information about the threshold used to ascertain the MID

Strength of the current work

- Include large proportion of the available MID literature
- Rigorous review methods including
  - Duplicate
  - Independent identification and selection of studies
  - Data extraction and reporting
  - Credibility assessment →use of a new reliable instrument to evaluate the credibility of MIDs

Limitation of current work

- No available reporting standard for anchor-based MIDs

- In the absence of such a guideline
  - Review of the relevant literature, required considerable judgment
  - Others may have chosen different criteria

Impact of reporting issues on credibility assessment and MID selection

- 13% of the studies did not report whether an intervention was used although determining the MID
- **Effective intervention**
- Increase the size of the difference in the PROM score between
  - Groups receiving and not receiving such an intervention
  - Between responders and non-responders
  - Influence the magnitude of the MID
- Surgical interventions may be associated with larger MIDs than nonsurgical
- Knowing the intervention
  - Allow MID users to select MIDs more effectively for PROMs

- PROM
  - 15-25% of reports, include lack of reporting of the lower and upper values that the PROM can reach, and the meaning associated those values

- MID estimation issues
  - 65% present only point estimate for MIDs with out variability (95%CI)
  - Ignoring variability and considering only point estimates → create a false sense of inconsistency across different MIDs for the same PROM
  - Users choosing among different MIDs for the same PROM → Prefer the one with more precise estimate

Correlation between anchor and PROM

- Anchor
  - External criteria to inform PROM interpretability
  - Correlation between anchor and PROM represents relatedness of the constructs
  - At least a moderate correlation, an MID is at best suspect
- Correlation between PROM and anchor and the precision of the MID estimation are linked
- Higher correlations can be associated with more precise MID estimates
- 66% of situations, authors failed to report correlation between anchor and PROM

Threshold to determine the MID estimate

- Threshold or a difference between groups in relation to an external criterion that represents a small but important difference

- Failure of MID estimation studies to define a threshold or use of a threshold that does not reflect a small but important difference undermines their credibility

- 16% of MIDs, authors failed to report the threshold associated with the MID, leading to a classification of ‘‘impossible to tell”

Recommendations to improve the reporting of MID estimation studies

- 1) Appropriately describing the settings
- 2) Describing the intervention used in the study → Allow users to compare across different MID studies
- 3) Providing a detailed of the PROMs they studied, including, range of values associated with the measure along with the meaning for the extreme values
- 4) Providing the number of participants and the threshold or difference between groups that was chosen to represent the MID
- 5) Reporting measures of variability accompanying the MID point estimate,
- 6) Measuring and reporting the correlation between the anchor and the PROM to which an MID is estimated

Conclusion

- Serious issues of incomplete reporting in the MID literature
  - Threaten the optimal use of MID estimates to inform the magnitude of effects of interventions on PROMs.
- Our suggestions
  - Should have a guideline for authors to report
- Systematically developed, consensus-based reporting checklist would help to achieve high reporting standards in the MID literature

Thank you for your attention