INTERPRETATION OF CLINICAL SIGNIFICANCE IN PSYCHOLOGICAL RESEARCH

Naomi M. Gutierrez, Laura Mills, Debra J. Pepler

YORK
UNIVERSITÉ
UNIVERSITY

INTRODUCTION

- A common question asked of administration, executive, and clinical teams is, "*What is the success rate of this program*?"
  - If the interest is in client health improvement, the question more specifically is, 'Do clients improve in measures of health during treatment and are treatment gains sustained for some time after the service?'
  - Fundamental to this question, and central to this research, is the issue of **significant**, or **meaningful clinical change**.

CLINICAL SIGNIFICANCE

- Addresses the practical importance of an effect, wherein said effect may represent a change at a group- and/or individual level.
- Suggests improvement and/or recovery(Ogles et al., 2001) or deterioration.
- A significant *p*-value (i.e., less than the typical criterion .05) is not evidence for or against clinical significance.
- The presence of a large effect size, whether standardized or not, does not mean that an effect is clinically significant.

GROUP-LEVEL MEASURES

Equivalence Testing

- Measures the recovery aspect of clinical significance (Kendall et al., 1999)

A clinical researcher chooses an **equivalence interval** of **[-0.58, 0.58]** in a mean difference context and obtains a **100(1-2α)% CI of [-0.23, 0.65]**. Since **the CI does not fall within the EI**, the researcher would not conclude that the means between the clinical and non-clinical samples are equivalent (i.e., **would not conclude clinical recovery**).

INDIVIDUAL-LEVEL MEASURES

Reliable Change Index

Measures Improvement (or deterioration) using change scores (i.e., post-treatment and pre-treatment difference scores) to determine whether individual change was reliable (i.e., not due to noise) (Jacobson & Truax, 1991).

$$RCI = \frac{x2 - x1}{\sqrt{2\left(SE^2\right)}}$$

- **References to normal functioning**

Three methods to calculate a cut-off score for a recovery/unchanged assessment based on the availability of normative data.

INDIVIDUAL-LEVEL MEASURES

Reliable Change Index

	Cut-off	Reliable change
Recovered	✓	✓
Improved	✗	✓
Unchanged	✗	✗
Deteriorated	✗	✓

GROUP- & INDIVIDUAL-LEVEL

Cut-off scores

Point-in-time scores (e.g., post-treatment scores)used as thresholds to assess recovery.

Elkin et al. (1989) defined recovery at any given time in terms of a cut-off score of **six or less** in the Hamilton Rating scale for Depression and a cut-off score of **nine or less** on the Beck Depression Inventory.

Percent Improvement (PI)

Compares pre-treatment scores to post-treatment scores to express improvement as a percentage (Blanchard & Schwarz, 1988).

GROUP- & INDIVIDUAL-LEVEL

Minimal Clinically Important Change Score (MCID)

The smallest change in magnitude of the measuring instrument that individuals perceive as important (whether beneficial or harmful).

There is no standard method for its calculation (Cook, 2008).

Franceschini et al. (2023) calculated and compared the MCID for knee osteoarthritis patients. The percentage of patients reaching the MCID varied based on the calculation method, with **anchor-based methods** showing percentages from **24.0% to 66.0%** and **distribution-based methods** from **44.6% to 75.9%**.

CODING

- We conducted a systematic literature review using PsycInfo and Google Scholar.

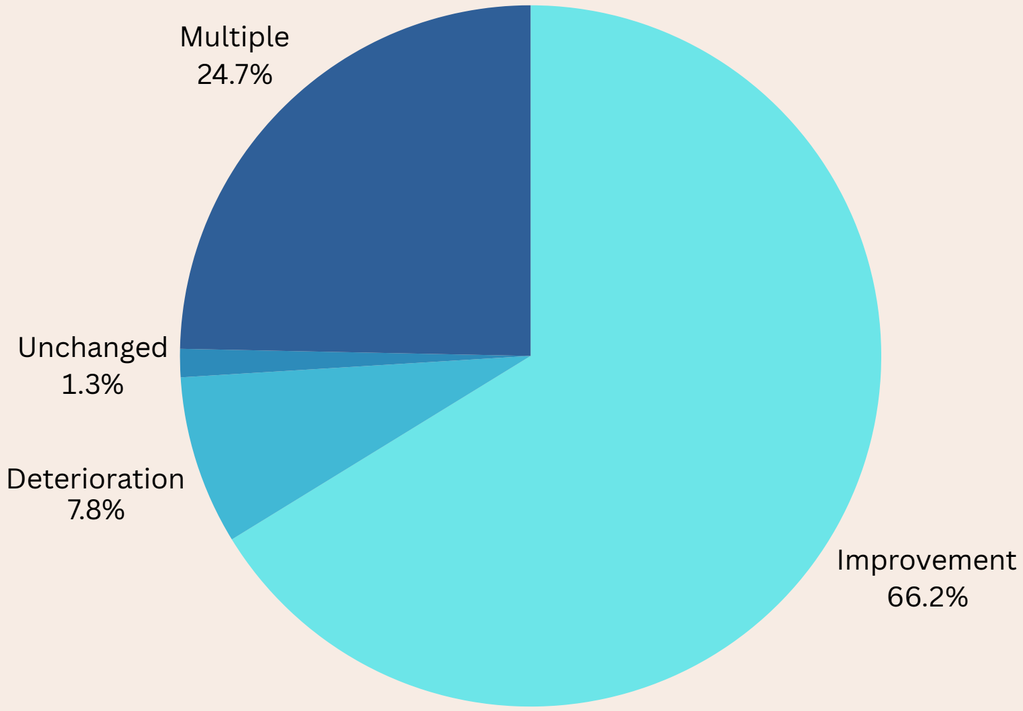
  Restricted our searches to the year 2023.

- Included empirical articles that contained at least one of the three terms in the results and/or discussion: **clinically significant**, **clinically meaningful** and **clinically important**.

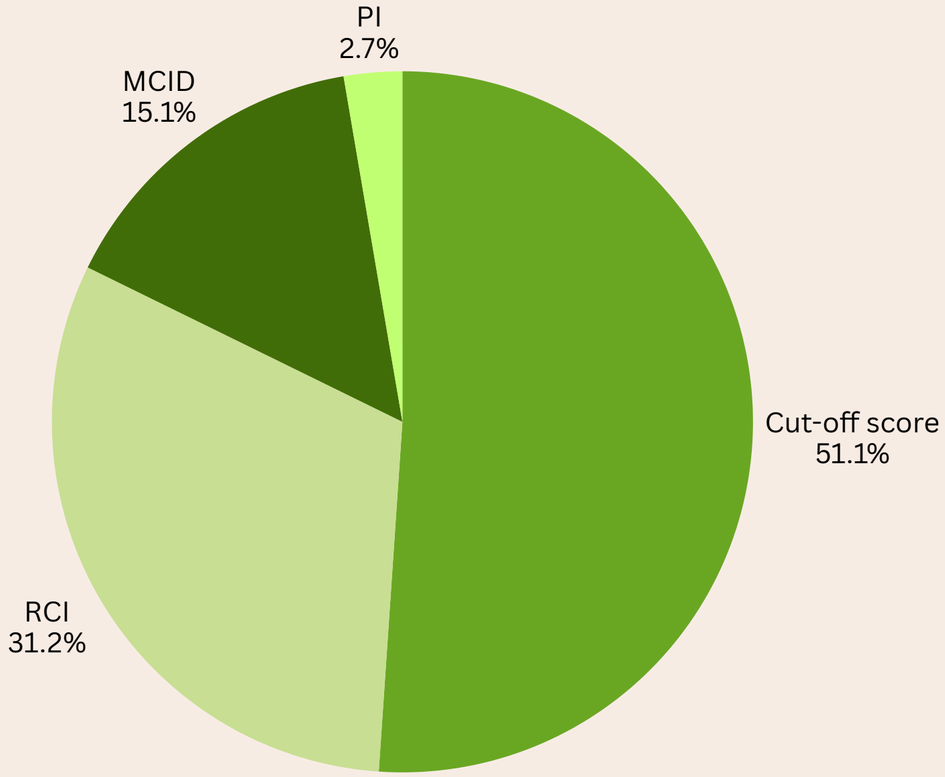
- Coded variables: **Clinical Significance Measure**; **Nature of Analysis**; **Focus**.

  If Clinical Significance Measure was an RCI, further coded on whether the RCI was instrument- or individual-based; and if a reliability estimate was reported.

FOCUS (ARTICLE-LEVEL)



METHOD (ANALYSIS-LEVEL)



RCI analyses

- 47% (n = 27) accompanied by a cut-off score.

- Use of instrument-based RCIs (n = 30) versus individual RCIs (n = 24) was approximately split (52% and 41%, respectively).

- Most(n = 47; 81%) did not explicitly define the type of reliability estimate incorporated to calculate the RCI.

  Of those that did, Cronbach's alpha was the most common choice (n = 10; 91%).

- Reported in a descriptive manner (n = 28; 48%) or as an outcome (n = 29; 50%).

CONCLUSIONS

- The discussion of clinical significance largely focuses on one direction of change.

  We therefore recommend to consider not only one direction of change (improvement), but also deterioration and stagnation.

- If RCI is the best measure of clinical significance, we recommend explicit reporting of the reliability estimate used in its calculation.

  Consider using use test-retest estimates instead of internal consistency measures (e.g., Cronbach alpha).

CONCLUSIONS

- We advise a cut-off for a non-clinical sample to be used alongside an RCI to conclude clinical significance.

- Re-consider using RCI scores as an outcome due to the potential loss of information.

THANK YOU

naomimg@yorku.ca

laura.mills@bestnotes.com