DEPARTMENT OF SPORTS SCIENCE AND CLINICAL BIOMECHANICS

SDU

Interpretation

Vodcast 3

$2 > -3$ ∞ $+ - \times \div$ $0.999... = 1$ $\pi \approx 3.14$ $\sqrt{2}$ 5^2 $1 + 2 \cdot 3$ $(1 - 2) + 3$ $5(2 + 2)$ $101_2 = 5_{10}$

Henrik Hein Lauridsen
Associate Professor
Research unit for Clinical Biomechanics
University of Southern Denmark

1

Questions

Responsiveness vodcast

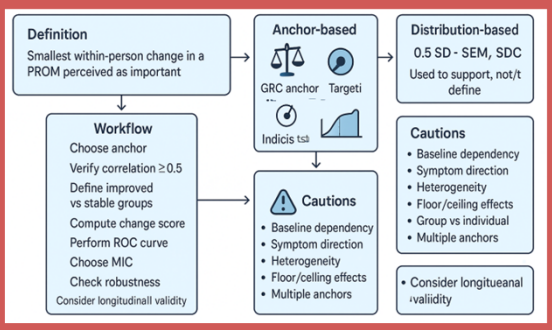
MIC video podcast I and II



3

Overview

- MIC credibility
- Linking measurement error to MIC
- Responder analysis
- Floor and ceiling effects



SDU

4

MIC credibility

Can we trust the anchor/transition rating?
Can we trust the MIC?

MIC credibility

"The extent to which the design and conduct of studies measuring MICs is likely to have protected against misleading estimates"

Carrasco-Labra A et al. Minimal important difference estimates for patient-reported outcomes: A systematic survey. J Clin Epidemiol. 2021



DEPARTMENT OF SPORTS SCIENCE AND

5

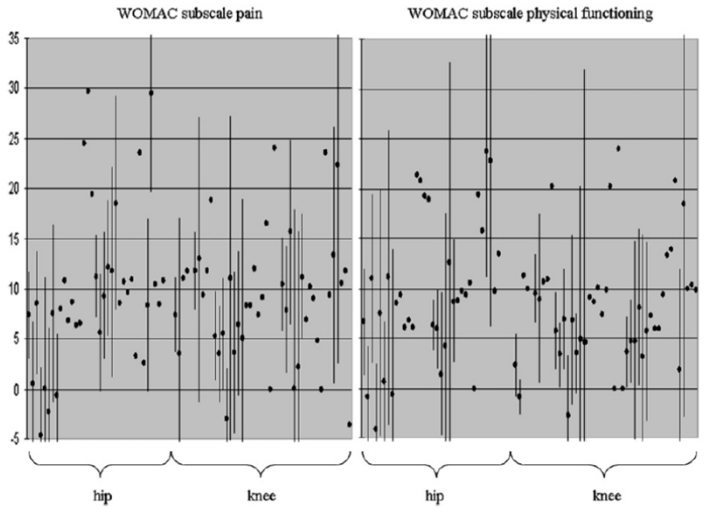
MIC variation

MIC varies according to

- Disease under scrutiny
- Disease severity
- Length of follow-up
- Methodology used
- Statistical approach used → Probably the most important aspects

- Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol. 2010 May;63(5):524–34.
- Hao Q, Devji T, Zeraatkar D, et al, Minimal important differences for improvement in shoulder condition patient-reported outcomes: a systematic review to inform a BMJ Rapid Recommendation. BMJ Open, 2019
- Devji T, Guyatt GH, Lytvyn L, et al. Application of minimal important differences in degenerative knee disease outcomes: a systematic review and case study to inform BMJ Rapid Recommendations. BMJ Open 2017

WOMAC - Western Ontario and McMaster University Osteoarthritis Index



Terwee CB et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol. 2010

Credibility criteria for external anchors

Development of 5 criteria for evaluating the credibility of an anchor

- Litterature review evaluating the completeness of reporting on anchor based MID's
- Expert steering committee

Two parts to the credibility instrument

- 5 core criteria
- 4 additional criteria on transition rating anchors



- Carrasco-Labra A, Devji T, Qasim A, Phillips MR, Wang Y, Johnston BC, et al. Minimal important difference estimates for patient-reported outcomes: A systematic survey. J Clin Epidemiol. 2021 May;133:61–71.
- Devji T, Carrasco-Labra A, Qasim A, Phillips M, Johnston BC, Devasenapathy N, et al. Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. BMJ. 2020 Jun 4;369:m1714.

Core criteria

- Is the patient or necessary proxy responding directly to both the PROM and the anchor?
  - Is it purely patient-rated or clinician rated?
- Is the anchor easily understandable and relevant for patients or necessary proxy?
  - Is the patient able to understand the anchor and use it easily for decision-making?
- Has the anchor shown good correlation with the PROM?
  - Good correlation between transition rating and PROM change score (0.5-0.7) → Terwee *et al.* (2021) recommends >0.3
- Is the MID precise?
  - Precision: [CI interval]/(MIC * 100) or sample size
  - Example: MIC = 23.5 [23.1; 23,8] → precision = (23.8-23.1)/23.5 * 100 = 3%
- Does the threshold or difference between groups on the anchor used to estimate the MID reflect a small but important difference?
  - Did the analyses calculate a small but important difference? How was the anchor used?

Example

CREDIBILITY ASSESSMENT – CORE CRITERIA

Is the patient or necessary proxy responding directly to both the PROM and the anchor?

If a clinician or anyone else is responding to the anchor directly and the patients are capable of providing this information, the answer should be "no." Any other necessary proxy (e.g. caregiver, parent, wife, relative) responding to the anchor, the answer is "yes".

[ ] Yes [ ] No [ ] Impossible to tell

Rationale:

Is the anchor easily understandable and relevant for patients or necessary proxy?

With "easily understandable and relevant" we mean that, when presented with the anchor (either actually presented or hypothetically) as an outcome, and without too much education, the patients are able to understand the data provided for the outcome (anchor) and use it easily for decision-making. For example, when addressing a multi-item PROM addressing the potential therapeutic effects of an intervention for iron-deficiency anemia, an anchor of patient's global rating of improvement in fatigue may be easier to understand and more relevant for decision-making than serum iron levels. If you were a patient, how would you answer this question?

[ ] Definitely yes [ ] To a great extent [ ] Not so much [ ] Definitely no [ ] Impossible to tell

Rationale:

10

Additional criteria

- Is the amount of elapsed time between baseline and follow-up measurement for MID estimation optimal?
  - Recall bias

To do with Present State Bias

- Does the transition item have a satisfactory correlation with the PROM score at follow-up?
  - Should be 0.1-0.2 and + (if same direction)
- Does the transition item correlate with the PROM score at baseline?
  - Should preferably be <0.1 or even negative (if same direction)
- Is the correlation of the transition item with the PROM change score appreciably greater than the correlation of the transition item with the PROM score at follow-up?
  - Difference should preferably be ≥ 0.1

11

Example

CREDIBILITY ASSESSMENT – ADDITIONAL CRITERIA FOR TRANSITION RATING ANCHORS

Is the amount of elapsed time between baseline and follow-up measurement for MID estimation optimal?

If there is a range of follow-up reported, consider the following when making your judgment: If the range falls over 3 categories (e.g. 3 weeks to 3 months), then select the middle category (i.e. in this example, you would select 'to a great extent'); If the range falls over 2 categories (e.g. 6 weeks to 3 months), then select the more conservative option (longest follow-up) (i.e. in this example, you would select 'not so much')

[ ] Definitely yes	[ ] To a great extent	[ ] Not so much	[ ] Definitely no	[ ] Impossible to tell
≤ 1 month	>1 to ≤2 months	>2 to ≤3 months	>3 months	

Rationale:

To answer the next 3 questions, you first need to determine if the scale of the anchor and PROM are in the same direction.

*For each question we provide 2 guides: If higher values on the anchor and PROM represent the SAME state (i.e. both represent a better or worse condition), use **Guide A**; If higher values on the anchor and PROM represent DIFFERENT states (i.e. higher scores on the PROM are worse, while higher values on the anchor are better), use **Guide B***

Does the transition item have a satisfactory correlation with the PROM score at follow-up?

Guide A

[ ] Definitely yes	[ ] To a great extent	[ ] Not so much	[ ] Definitely no	[ ] Impossible to tell
>0.2	0.1 to 0.2	<0.1	Negative correlation	

Guide B

[ ] Definitely yes	[ ] To a great extent	[ ] Not so much	[ ] Definitely no	[ ] Impossible to tell
<-0.2	-0.1 to -0.2	>-0.1	Positive correlation	

12

Items relevant for reporting/selecting MID

ELSEVIER

Journal of Clinical Epidemiology 142 (2022) 144–151

Journal of Clinical Epidemiology

REVIEW

A systematic survey identified methodological issues in studies estimating anchor-based minimal important differences in patient-reported outcomes

Yuting Wang[a,*], Tahira Devji[a], Anila Qasim[a], Qiukui Hao[a,b], Vanessa Wong[a], Meha Bhatt[a], Manya Prasad[c], Ying Wang[a], Atefeh Noori[a], Yingqi Xiao[a,d], Maryam Ghadimi[a], Luis Enrique Colunga Lozano[e], Mark R. Phillips[a], Alonso Carrasco-Labra[f], Madeleine King[g], Berend Terluin[h], Caroline B Terwee[i], Michael Walsh[a,j,k], Toshi A. Furukawa[l], Gordon H. Guyatt[a]

U.S. terminology

Table 2. List of identified Items relevant to reporting and/or selecting anchor-based MID*

Category	Code	Item	N‡
Definition	MID definition	MID definition †	95
	MID terminology	Standardize MID terminology †	6
Anchor issues	Anchor type	Possible anchors†	30
	Anchor perspective	The change being perceived as minimally important on the anchor would be perceived as beneficial from the patient's viewpoint.	30
		Proxy perspective is acceptable if patients cannot make decisions.	4
		Estimating MID should use clinical evaluation as anchor.	19
	Anchor threshold	The choice of what constitutes an MID on the anchor should reflect a small but important difference .	38
	Anchor interpretability	Anchor should be easily understandable.	18
	Correlation between PROM and anchor	At least moderate correlation between PROM and anchor.	36
		Transition scale should correlate equally strongly with the pretest and posttest PROM scores but with opposite signs.	6
		The correlation of transition scale with change scores on PROM should be at least 0.2 greater (in absolute terms) than the correlations with either the baseline or the follow-up test scores.	4
	Anchor construct	Anchor and PROMs should measure the same or closely related constructs.	20
	Recall time of transition scale	The recall time should not be too long when using transition scale as anchor.	53
	Unchanged group of transition scale	The use of unchanged group when using transition scale as anchor.	5
	Number of anchors	Multiple anchors should be used to measure MID, rather than single anchor.	10
	Measurement properties of anchor	Satisfactory anchor measurement properties (validity, reliability and responsiveness) are the prerequisite of an optimal anchor.	35
PROM issue	Measurement properties of PROM	Required measurement properties for PROM when measuring MID: reliability, validity and responsiveness.	4
		Ordinal scales do not support the mathematical calculations required for MCID calculation.	5
Generalizability of MID	MID varying by contexts	MID may vary depending on patient characteristics.	86
		MID may vary depending on follow-up length.	15
		MID may vary depending on intervention.	28
	MID direction	Distinguishing MID for improvement vs deterioration.	31
	Generic instruments	For generic instruments, MID should not vary depending on the contexts.	3
Statistical issues	Precision of MID	Precision of MID	19
	MID in relation to measurement error	MID should be beyond measurement error	16
	Triangulation	Triangulation of distribution-based and anchor-based methods.	15
	Representative sample	Representative sample.	14
	Relative or absolute MID	Expressing MID in relative or absolute terms †	19
	Analytical method	Possible anchor-based MID analytical methods †	78
		MID may vary depending on the analytical method used.	33
Others	MID selection	MID selection: match study context	8
		MID selection: evidence-based and consensus processes	6
		MID selection: MID levels	1
	A single MID or range of plausible MIDs	A single value or range of plausible MIDs is appropriate for many PROMs	29

Note: * see *Appendix 4* for the best example of quotations for items.
† Apply to anchor-based MID reporting only.
‡ Number of citations

13

Linking measurement error to MIC



14

MIC is conceptually different from SDC

SDC is the smallest change in score that you **CAN** detect with the instrument, above measurement error

MIC is the smallest change in score that you **WANT** to detect with the instrument

In individual patients SDC < MIC to distinguish important change from measurement error

Often this is not the case

15

An example

Qual Life Res (2008) 17:1257–1267
DOI 10.1007/s11136-008-9409-3

Comparative responsiveness and minimal change for the Oxford Elbow Score following surgery

Jill Dawson · Helen Doll · Irene Boller · Ray Fitzpatrick · Christopher Little · Jonathan Rees · Andrew Carr

Measurement error (SDC)

MIC

		ICC	Scaled SEM	Scaled MDC 90%	Best cut-point
Elbow pain					
OES					
Oxford Elbow Score	Function	0.89	8.23	18.73	5.0
	Pain	0.98	3.58	8.25	12.5
	Social-Psychological	0.87	8.51	18.85	12.5
	DASH	0.96[a]	4.06	9.33	−5.0

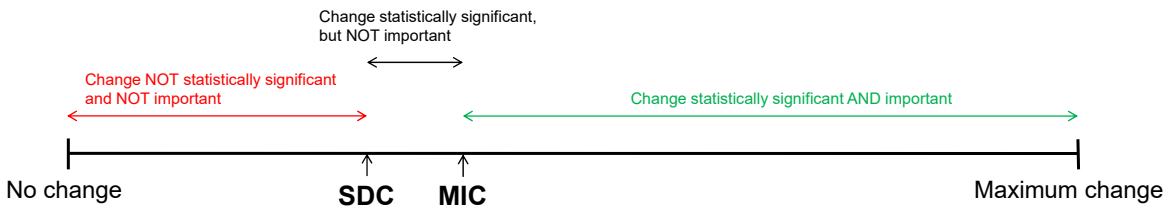
$OES_{function}$: SDC (19) is larger than MIC (5)
OES_{pain}: SDC (8) is smaller than MIC (12.5)

Therefore

OES_{pain} can distinguish important change from measurement error whereas $OES_{function}$ can not.

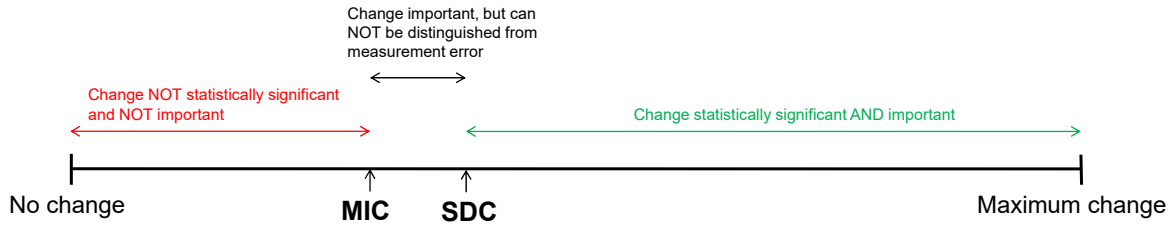
SDC is smaller than MIC

SDC and MIC are two reference points in the scale that can help interpret change scores



From example: OES_{pain} - SDC (8) is smaller than MIC (12.5)

SDC is larger than MIC



From example: $OES_{function}$ - SDC (19) is larger than MIC (5)

18

Taking type 1 error into account

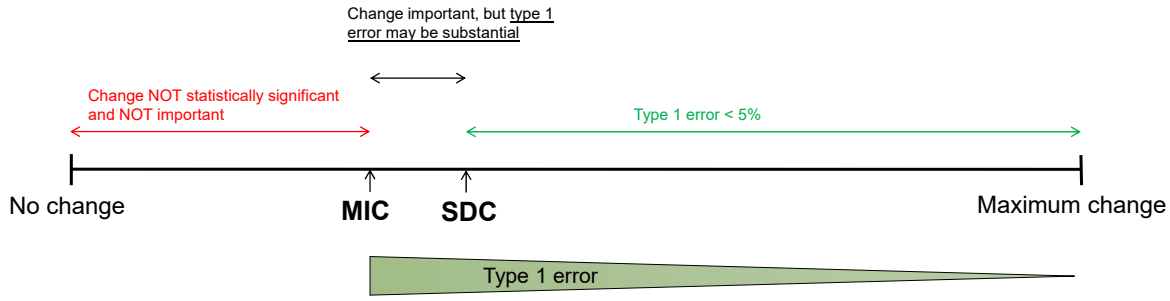
$$SDC_{agreement} = 1.96 \times \sqrt{2} \times SEM_{agreement}$$

- A patient who changed as much as the SDC has 'really' changed (statistically significant changed). However, there is a 5% probability (type 1 error) that in fact this patient has not changed
- Patients with change scores smaller than the SDC have a higher probability that they are in fact not changed (larger type 1 error)
- Thus, if the SDC is ***larger*** than the MIC, there is a higher type 1 error if we call patients who changed as much as the MIC importantly improved

19

Taking type 1 error into account

SDC is larger than MIC



20

How to use the MIC

In research

- Responder analysis

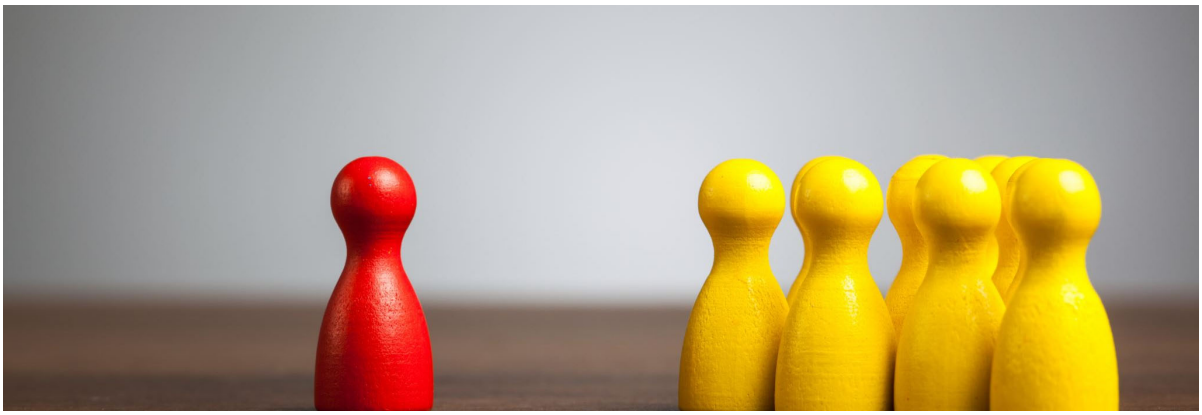
In clinical practice

- An example



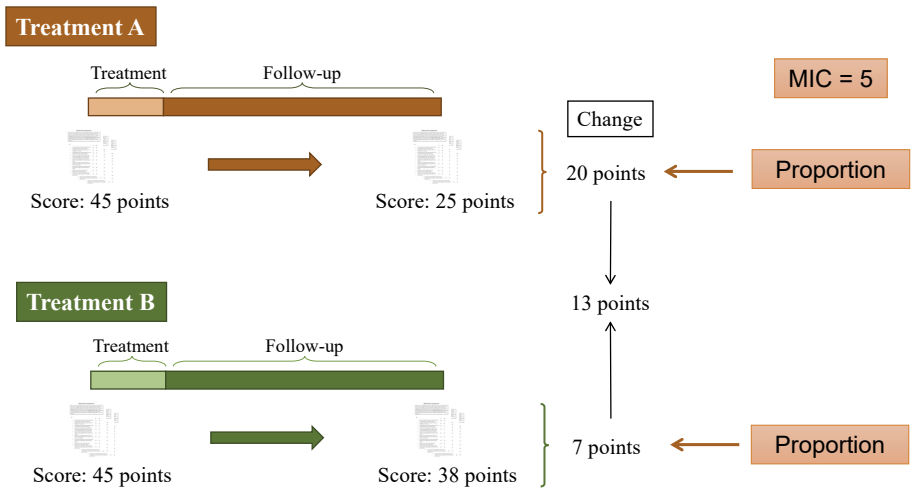
21

MIC and responder analysis



22

MIC and RCT's



23

NNT – dichotomous outcome

	Treatment	
Control	*Better (0.48) x*	*Unchanged (0.52) y*
Better (0.28) a	0.13 (ax)	0.15 (ay)
Unchanged (0.72) b	0.35 (bx)	0.37 (by)

Better in the treatment group Better in the control group

$$NNT = \frac{1}{bx - ay}$$

$$NNT = \frac{1}{0.35 - 0.15} = 5$$

Guyatt et al. (1998, 2002)

NNT – 3 outcomes

	Treatment		
Control	*Better (0.48) x*	*Unchanged (0.42) y*	*Worse (0.1) z*
Better (0.28) a	0.13 (ax)	0.12 (ay)	0.03 (az)
Unchanged (0.49) b	0.24 (bx)	0.21 (by)	0.05 (bz)
Worse (0.23) c	0.11 (cx)	0.10 (cy)	0.02 (cz)

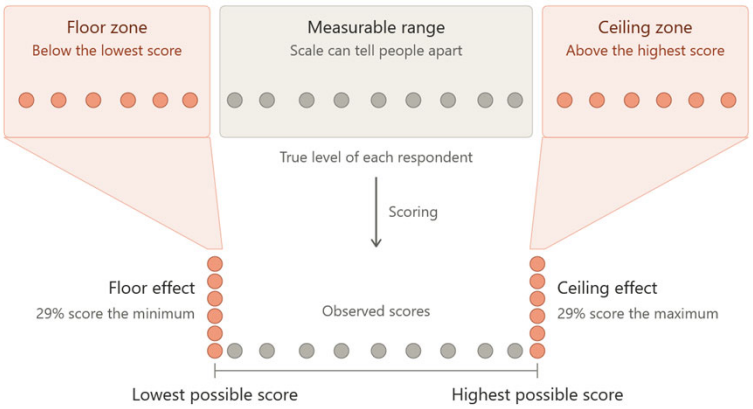
Better in the treatment group Better in the control group

$$NNT = \frac{1}{(bx + cx + cy) - (ay + az + bz)}$$

$$NNT = \frac{1}{(0.24 + 0.11 + 0.10) - (0.12 + 0.03 + 0.05)} = 4.2$$

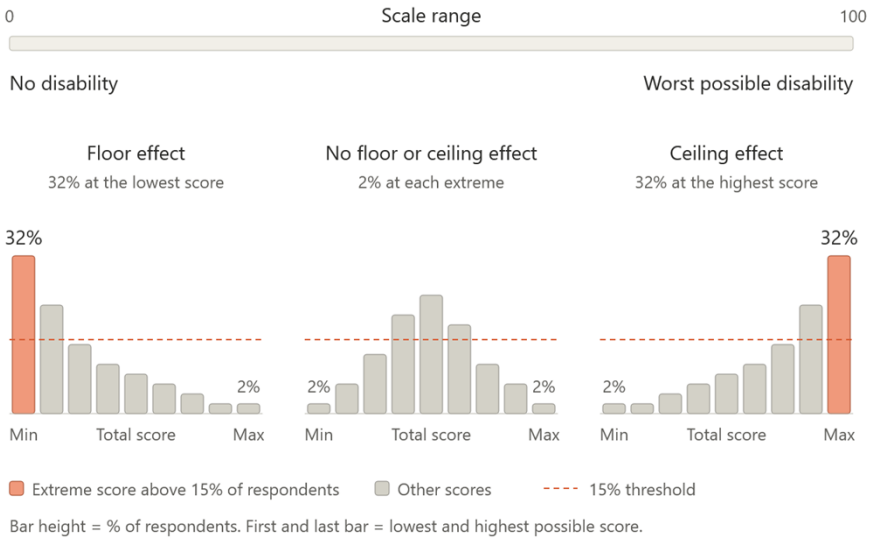
Guyatt et al. (1998, 2002)

Floor and ceiling effects



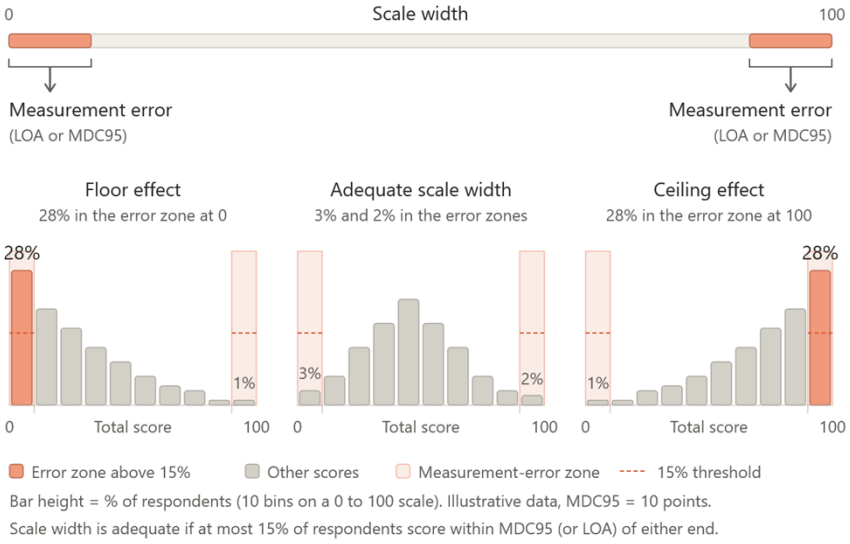
26

Floor and ceiling effects



27

Floor and ceiling effects



Davidson et al. 2002

28

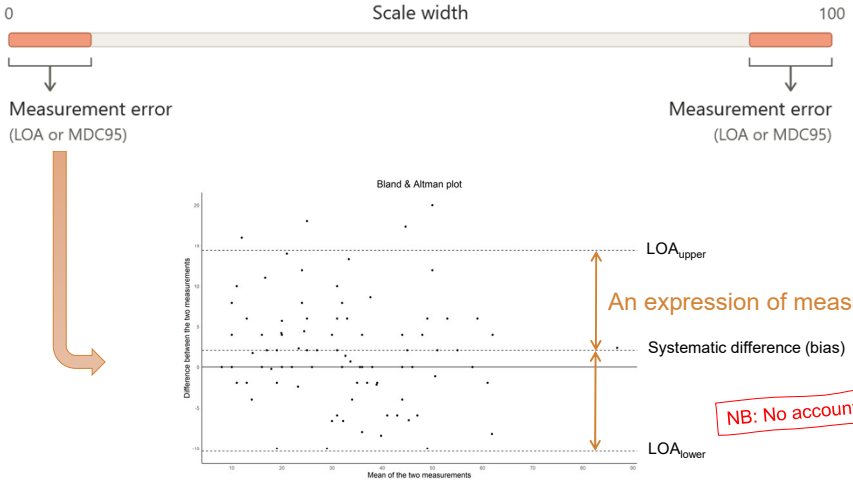
Example



	N	Oswestry Disability Index (%)
Floor effect	147	0 (0.0)
Ceiling effect	147	0 (0.0)

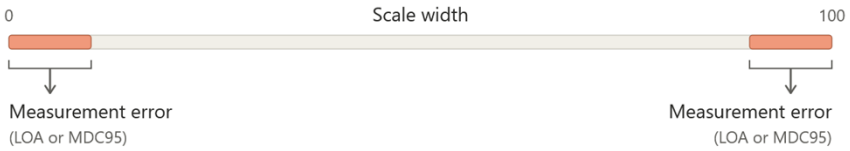
29

Floor and ceiling effect



30

Floor and ceiling effect



	N	Oswestry Disability Index (%)
Floor effect	147	9 (6.1)
Ceiling effect	147	1 (0.7)

31


THANK YOU FOR LISTENING