MEDIZINISCHE UNIVERSITÄT
INNSBRUCK

Minimal important differences (MID) for individual subjects and groups of subjects

Georg Kemmler, Johannes Giesinger,
Bernhard Holzner

Department of Psychiatry and Psychotherapy, Innsbruck Medical University
Innsbruck, Austria

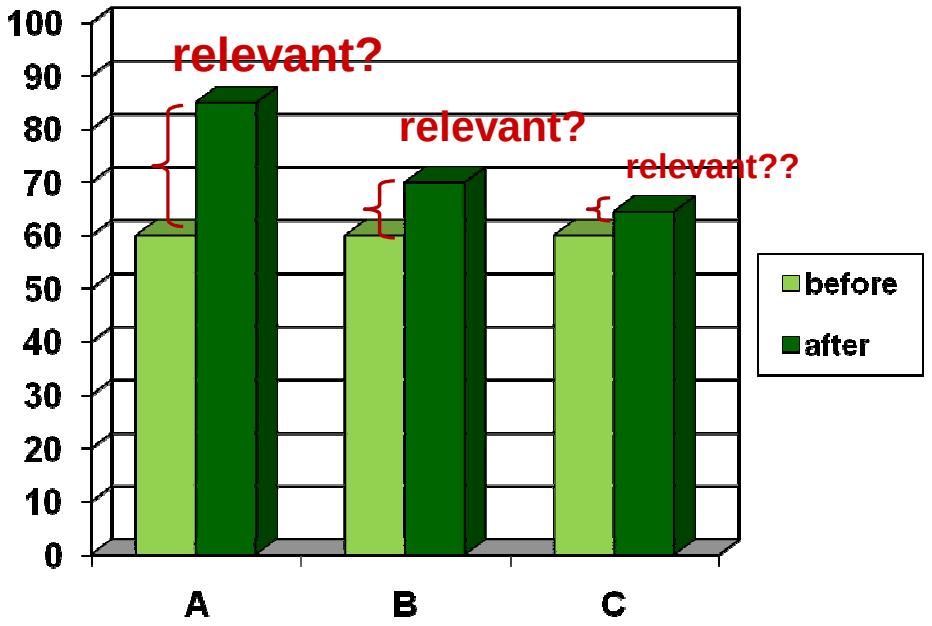
Overview

1. Background and research aims
2. Rational for a distinction between individual-level and group-level MID
3. Group-level MID based on the “number needed to treat”
4. Examples
5. Discussion

1 Background and research aims

- What is the Minimal Important Difference (MID)?

  The smallest difference or change on a scale considered clinically relevant



- How can the MID be determined?

  - distribution-based methods (e.g., MID = 0.5 * SD)
  - **anchor-based methods**

Anchor-based determination of the MID 1:

Examples of anchors

- **Patient self-rating of change in QOL, e.g.:**

  *How would you rate your actual quality of life compared to the last assessment? **

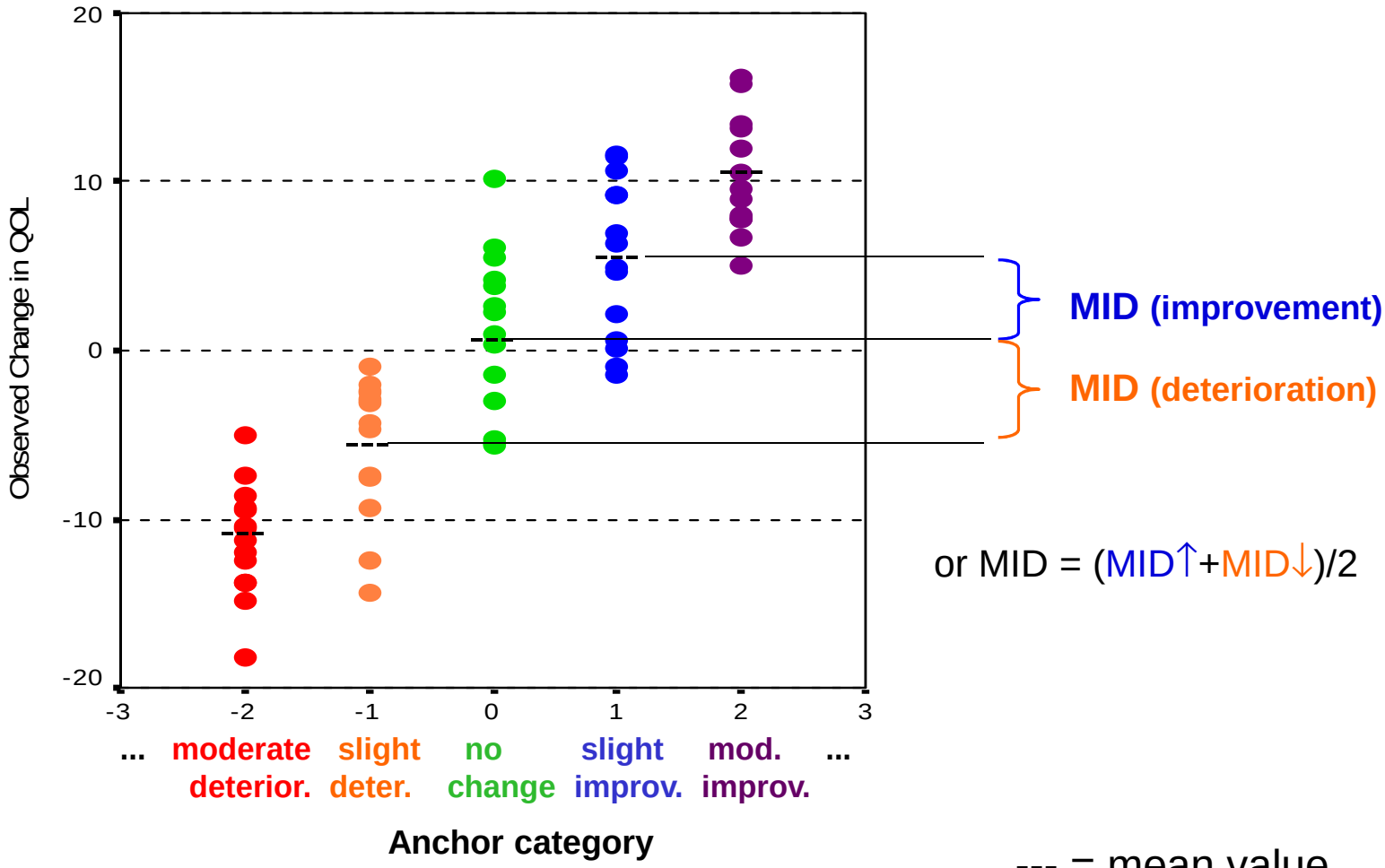
Much worse	*Moderately worse*	*A little worse*	*About the same*	*A little better*	*Moderately better*	*Much better*
-3	*-2*	*-1*	*0*	*1*	*2*	*3*

- Assessment by the doctor (similar as above)
- Laboratory parameter or other surrogate marker of health status
- …

* Osoba et al., J Clin. Oncol. 1998

Anchor-based determination of the MID 2:

Standard method for MID estimation



MIDs derived by this method

- EORTC QLQ-C30 subscales (cutoff values for changes)
  - 5-10 small change (=MID?)
  - 10-20 moderate change
  - > 20 large change

  Osoba et al. JCO 1998, Rodrigues et al. Qual Life Res 2004

- FACT-G and a number of FACIT modules

  Cella et al. 2002, Eton et al. 2004, Yost et al. 2005…
- SF-36
- …

Aims of this Talk

Usually **no distinction** is made between

- MIDs for **individual subjects** (e.g. indiv. patient monitoring) and
- MIDs for **groups of subjects** (e.g. treatment A vs. B in a clinical trial)

Aims

1. **Rational for a distinction between the two cases** (group, individual)
2. **Informal definition of a group-level MID** based on the "number needed to treat"
3. **Comparison of individual-level MID and group-level MID** for two examples

2. Rational for a distinction between individual- and group-level MID

- **Individual level**: focus on change in the **individual patient**
  - ➢ Anchor-based approach for MID is in principal OK
  - ➢ But: insufficient reliability/validity of the anchor leads to an underestimation of the MID*

- **Group level:** focus on one or several **groups of subjects**

  e.g. clinical trial with QOL as primary/secondary endpoint
  - ➢ MID for **mean change** in a group
  - ➢ Conclusions for individual subjects can not directly be drawn

* Yost et al., Eval Health Prof 2005; Kemmler et al. ISOQOL 2009

Example : one group, two assessment times

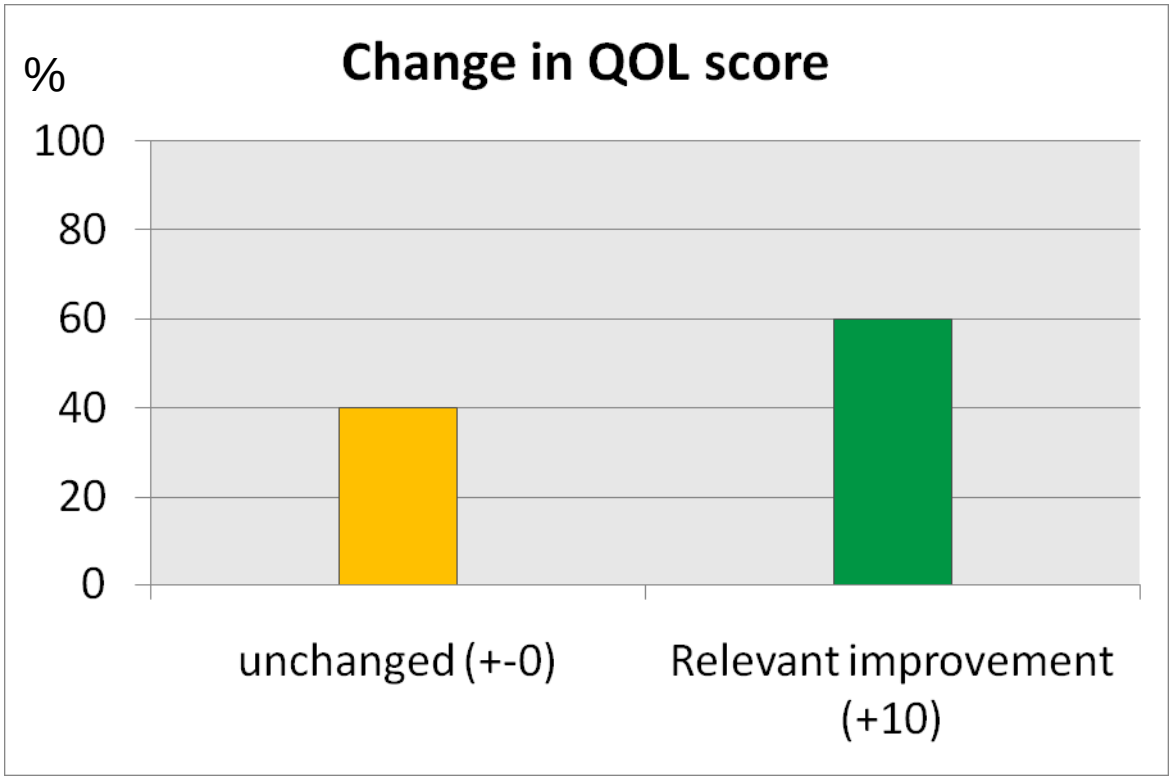
Suppose:

- Scale from 0-100
- MID (individual subject) = 10 points
- **60%** of subjects **show relevant improvement (+10 points)**
  40% of subjects remain **unchanged** (+- 0)

Then:

- Majority of subjects has improved!
- Mean change = +6 points (< 10 points)

Example (continued)



Mean change = +6

What happens, if we use MID=10 as threshold for relevant change in the group?

Change within group is considered irrelevant! (Mean change = +6 < +10 !)

Problem can be solved by introducing "group level MID"

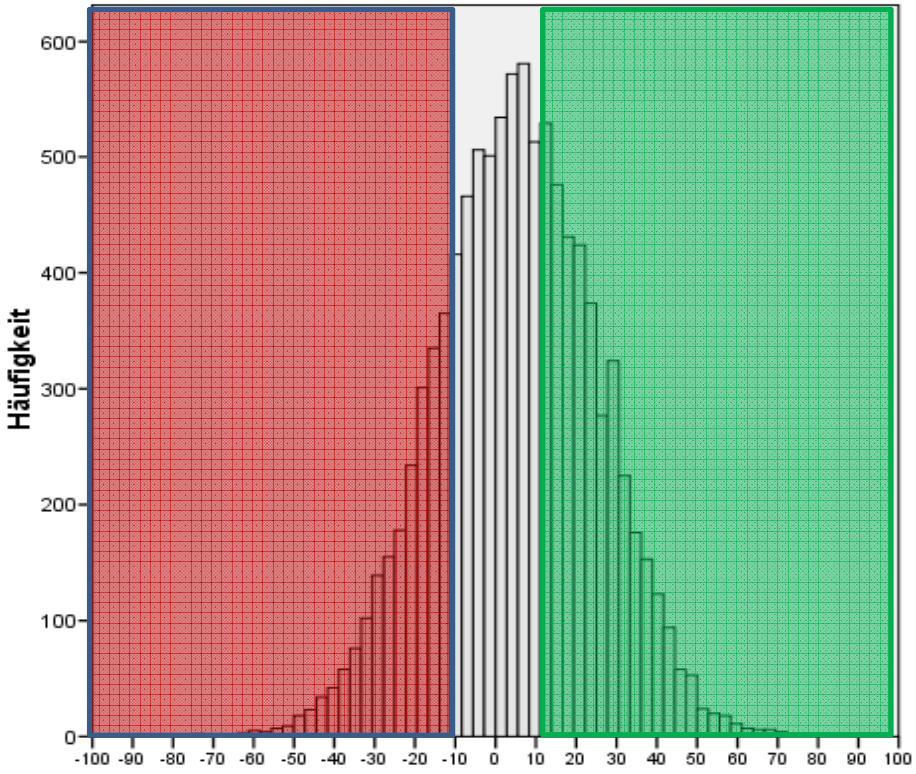
3. Group-level MID based on the "number needed to treat"

Given: A distribution of changes (real data or simulated data)

Calculate: **1) % subjects with relevant improvement** (based on indiv. level MID)

2) % subjects with relevant deterioration (" " " " ")

3) difference: % improved - % deteriorated



e.g.

45% improved
25% deteriorated

difference = 20%

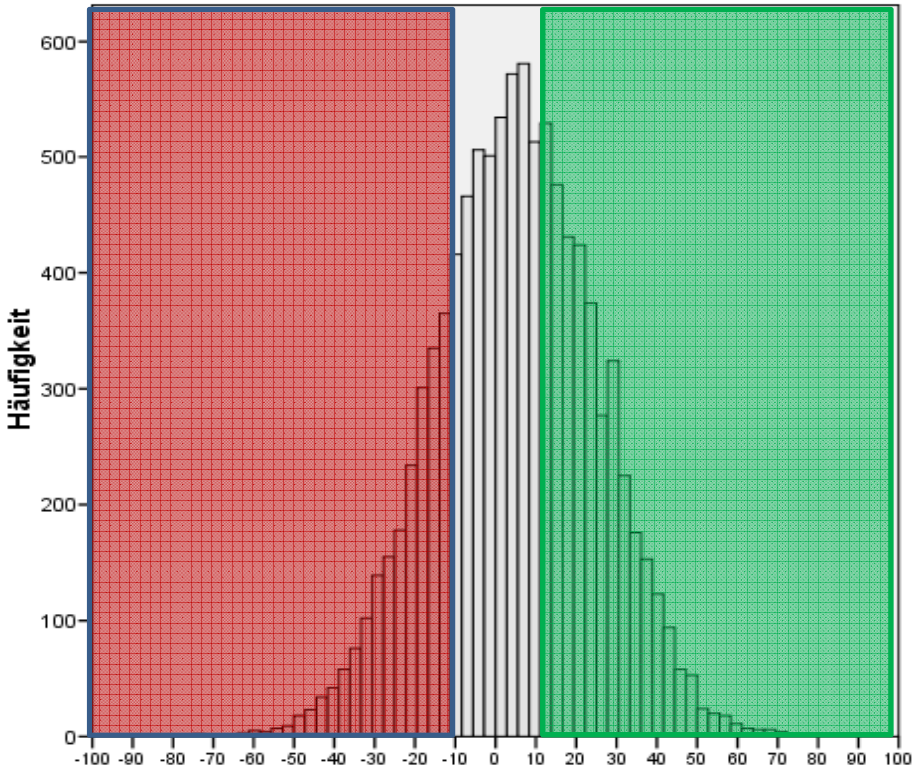
The “number needed to treat“

(framework for MID on the group level)

NNT: Number of subjects needed to treat *for **one** person to benefit*

(i.e. one subject **more** with improvement than with deterioration)

Formula: NNT = 1 / { P(improvement) – P(deterioration) }



e.g.

45% improved
25% deteriorated

difference = 20%

NNT = 1 / 0.2
= 5 subjects

NNT – threshold for „relevant benefit“?

- Literature does not give much guidance reg. cutoff for relevant NNT
- Wide range of NNTs in clinical examples: ≈ 3 – ≈ 30
- “Ad hoc cutoffs“:
  NNT = 5 (20% difference) rather too large(?)
  NNT = 10 (10% difference)
  NNT = 20 (5% difference) too small?

Provisional “definition“ of group-level MID

Suppose, a **mean change d** on a scale is associated with

10% (xy%) **more improvements than deteriorations** (indiv. level)

Then **d : = group-level MID**

4. Examples

Example 1: Pre-terminal cancer patients*

- N = 33
- various cancer diagnoses
- mean age 64.4 ± 11.1 (mean±SD) , 54% female

Two assessment times (selected from more assessments)

t1: 108 ± 64 d before †

t2: 25 ± 21 d before †

* Data provided by Natters State Hospital, Tyrol, Austria (Prof. H. Denz, Prof. M. Fiegl)

Example continued

Subscale	Mean t1	Mean t2	Mean change
PF	28.5	12.7	-15.8*
EF	53.3	41.7	-11.6*
SF	86.6	83.3	- 3.3
Pain	52.0	64.1	+12.1(*)
Fatigue	76.8	86.5	+9.7*
Taste alt.	54.0	61.6	+7.6
Diarrhea	18.1	19.1	+1.0

* $p < 0.05$, (*) $p < 0.09$, Wilcoxon matched-pairs test

Example continued

Subscale	Mean t1	Mean t2	|Mean change|	Relevant deteriora-tion (≤-10)	Relevant improve-ment (≥10)	Difference	NNT
PF	28.5	12.7	15.8				
Pain	52.0	64.1	12.1				
EF	53.3	41.7	11.6				
Fatigue	76.8	86.5	9.7				
Taste alt.	54.0	61.6	7.6				
SF	86.6	83.3	3.3				
Diarrhea	18.1	19.1	1.0				

Example continued

Subscale	Mean t1	Mean t2	|Mean change|	Relevant deterioration (≤-10)	Relevant improvement (≥10)	Difference	NNT*
PF	28.5	12.7	15.8	57.6%	0.0%	57.6%	<2
Pain	52.0	64.1	12.1	48.5%	21.2%	27.3%	4
EF	53.3	41.7	11.6	48.5%	30.3%	18.5%	6
Fatigue	76.8	86.5	9.7	51.5%	12.1%	39.4%	3
Taste alt.	54.0	61.6	7.6	40.0%	23.3%	16.7%	6
SF	86.6	83.3	3.3	13.3%	13.3%	0%	∞
Diarrhea	18.1	19.1	1.0	26.7%	23.3%	3.3%	30

⇒ Indication that group-level MID should be “somewhere between 3.3 and 7.6“

*Here: number needed to observe to see one ***more*** case with deterioration than with improvement

Example 2: Simulated Change Data

- QOL data often have skewed distribution
- But: ***changes*** have a more symmetrical shape
- SD of QLQ-C30 subscales: approx. 20 (+-)
- SD of changes of subscales: similar

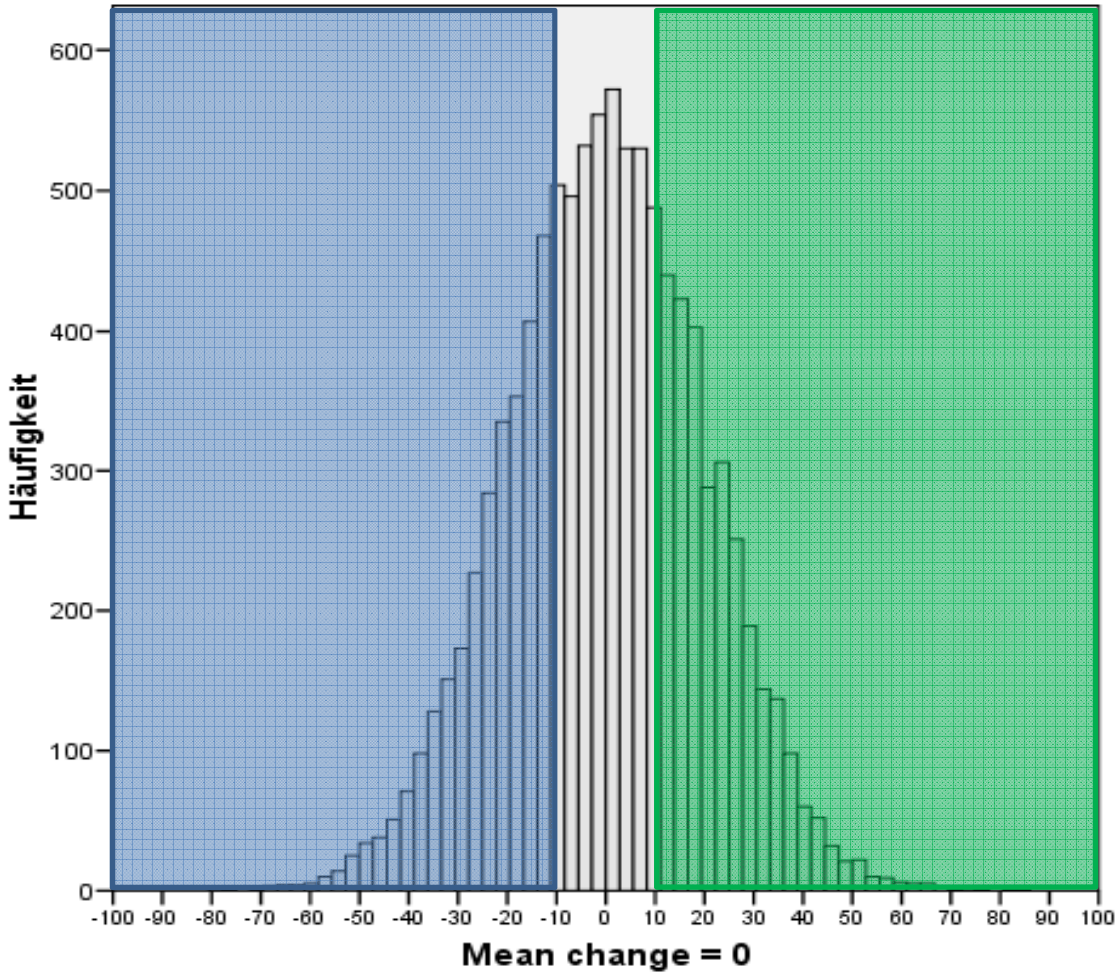
Therefore use following simulated changes:

- Normal distribution
- SD=20
- Let **mean change** vary: 0, 1, …, 10, …

Simulation of changes: Mean change = 0

Normal distribution, SD = 20, MID (individual subject) = 10

Relevant improvements and deteriorations



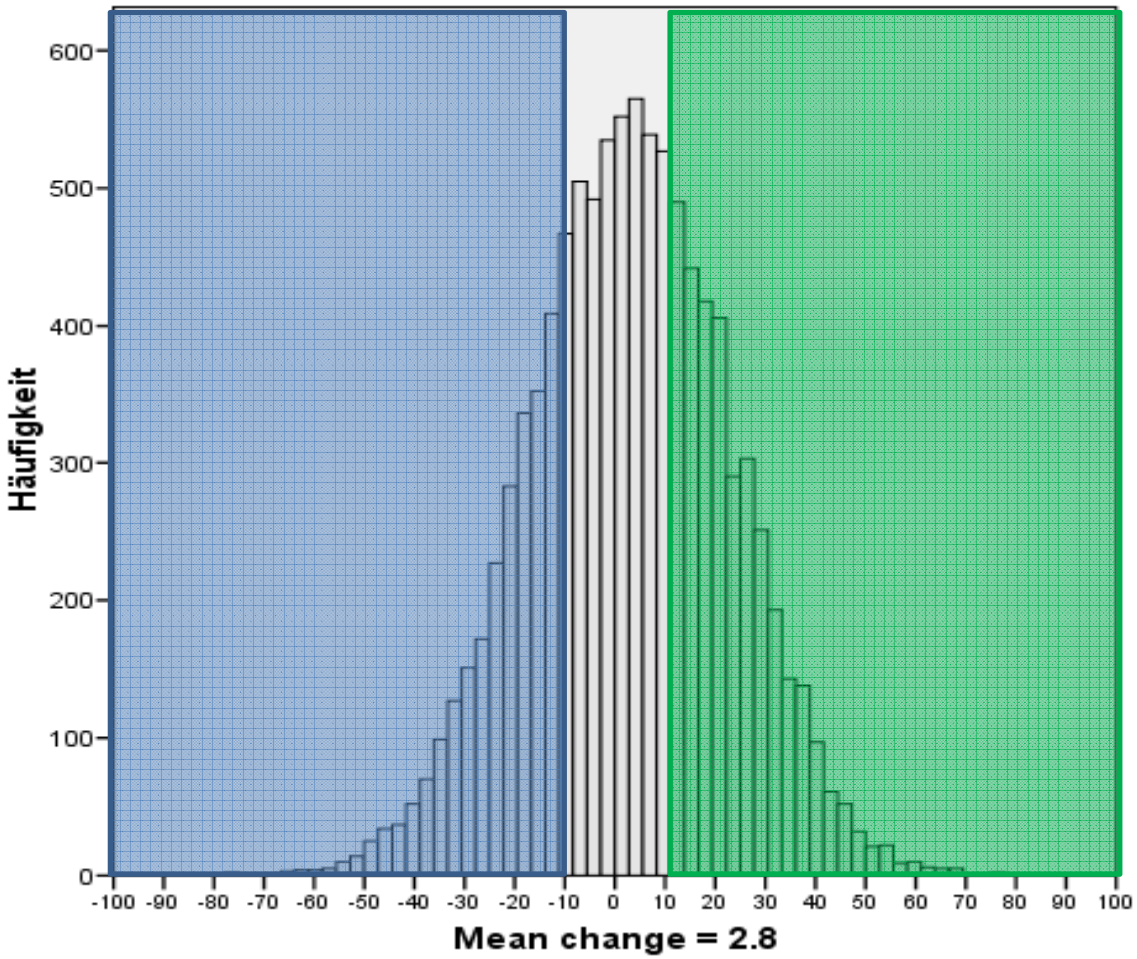
Relevant improvement (change > MID): **30.8%**

Relevant deterioration (change< -MID): **30.8%**

Simulation of changes: Mean change = 2.8

Normal distribution, SD = 20, MID (individual subject) = 10

Relevant improvements and deteriorations



Relevant improvement (change> MID): **36.0%**

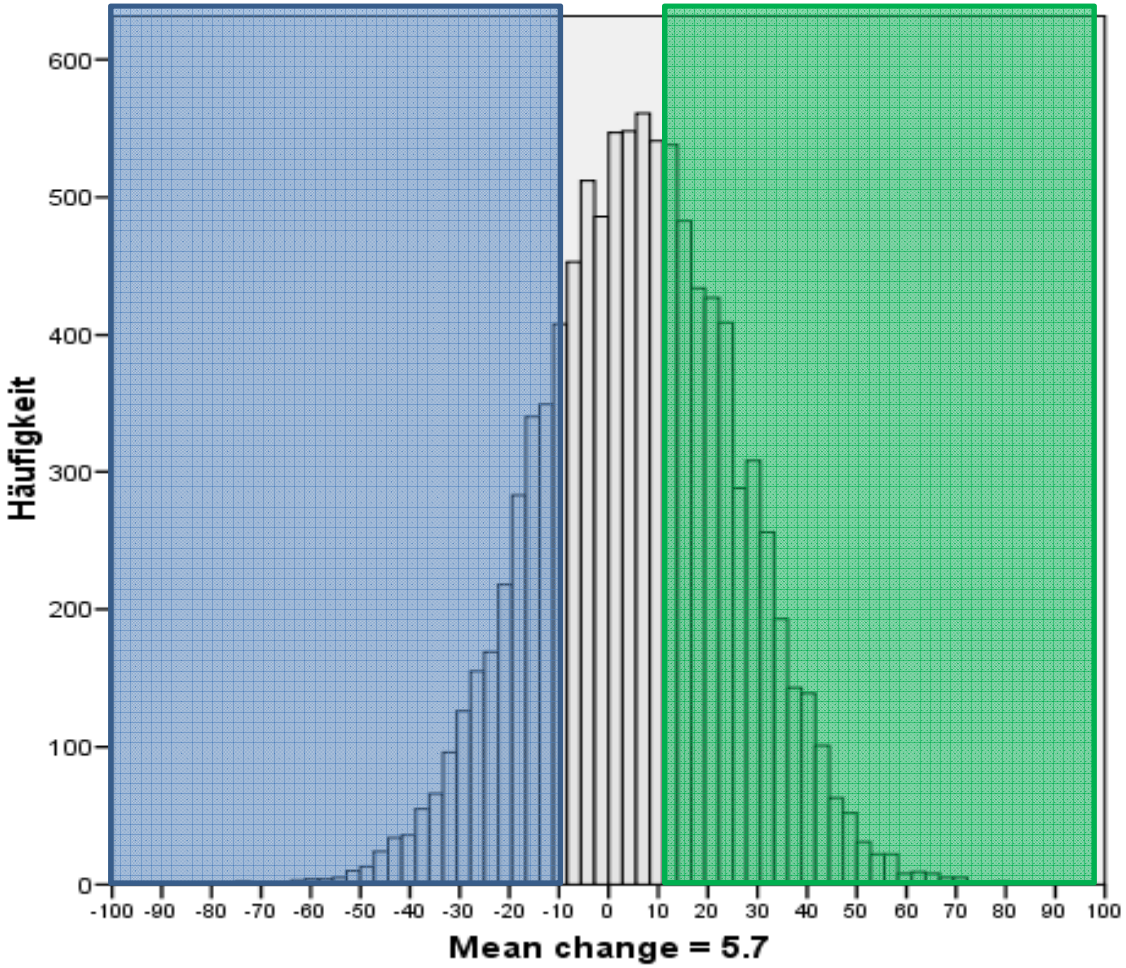
Relevant deterioration (change < -MID): **26.0%**

10% ***more*** improvements than deteriorations, NNT = 10

Simulation of changes: Mean change = 5.7

Normal distribution, SD = 20, MID = 10

Relevant improvements and deteriorations



Relevant improvement (change> MID): **41.5%**

Relevant deterioration (change < -MID): **21.5%**

Mittelwert =5,70
Std.-Abw. =20,00
N =10.000

20% ***more*** improvements than deteriorations, NNT = 5

Example 2 – Summary of results

MID individual level	SD of change	MID group level (NNT=10)	MID group level (NNT=5)
5	20	2.6	5.1
10	20	2.8	5.7
20	20	4.0	8.1
5	30	3.7	7.5
10	30	4.0	8.0
20	30	4.6	9.5

5. Discussion

- Distinction between group-level and individual level MID appears important (if reasoning here is correct)
- Analyses suggest:
  MID (group level) < MID (individual subject)
- Procedure to obtain group-level MID still has to be worked out in detail

* * *

- Is there need to revise individual-level MID? (larger thresholds?)
- Need to rethink the MID concept as a whole?