



Gütekriterien für evaluative Messinstrumente in der Rehabilitation

Dipl.-Psych. Wilmar Igl

*Rehabilitationswissenschaftlicher
Forschungsverbund Bayern (RFB)*

Universität Würzburg

Gütekriterien für Messinstrumente

- **Anwendungsziele** von klinischen Messinstrumenten (vgl. Kirshner & Guyatt, 1985):
 - Diskrimination: Messung stabiler Konstrukte z.B. Intelligenz
 - Prädiktion: Vorhersage von Werten z.B. Mortalität
 - Evaluation: Beurteilung von Veränderungen z. B. Reha-Erfolg
- Messinstrumente zur **Diskrimination** oder **Prädiktion**:
Klassische Gütekriterien (Objektivität, Reliabilität, Validität) und Kennwerte (z.B. Retest-Reliabilität) ausreichend
- Messinstrumente zur **Evaluation**:
Klassische Gütekriterien bzw. Kennwerte nicht ausreichend

⇒ „**Änderungssensitivität**“ **notwendig!**

Beispiel: Änderungssensitivität vs. Re-Test-Reliabilität

Fragebogen A	t1	t2	Intervention	t3	Differenz t3 – t2	Test - ergebnis
Vp1	8	9	EG	15	+6	++
Vp2	9	8	EG	15	+7	++
Vp3	8	9	KG	8	-1	=
Vp4	9	8	KG	9	+1	=
Fragebogen B	t1	t2	Intervention	t3	Differenz t3 – t2	Test - ergebnis
Vp1	5	5	EG	5	0	++
Vp2	9	9	EG	9	0	++
Vp3	5	5	KG	5	0	=
Vp4	9	9	KG	9	0	=

Tabelle 1: Beispiele für ein nicht reliables, aber änderungssensitives (A) bzw. ein reliables, aber nicht änderungssensitives Messinstrument (B) (nach Guyatt & al., 1987)

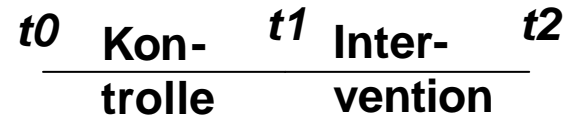
Änderungssensitivität – Messdesigns (1)

Ein-Gruppen-Designs zur Ermittlung der Änderungssensitivität

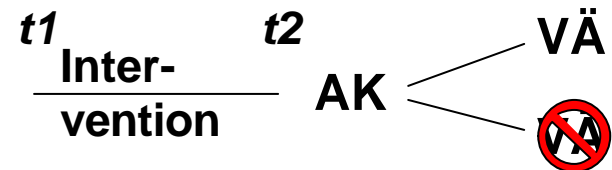
- Einfaches Prä-Post-Design



- Prä-Post-Design mit Baseline



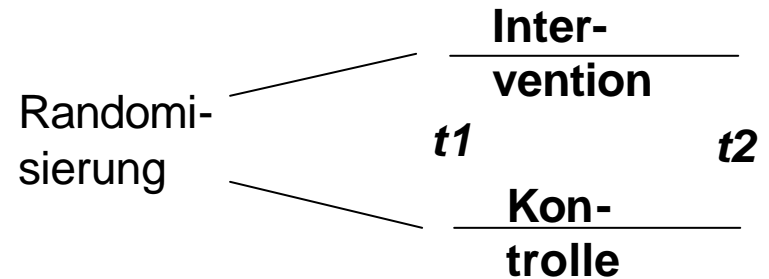
- Einfaches Prä-Post-Design mit Außenkriterium



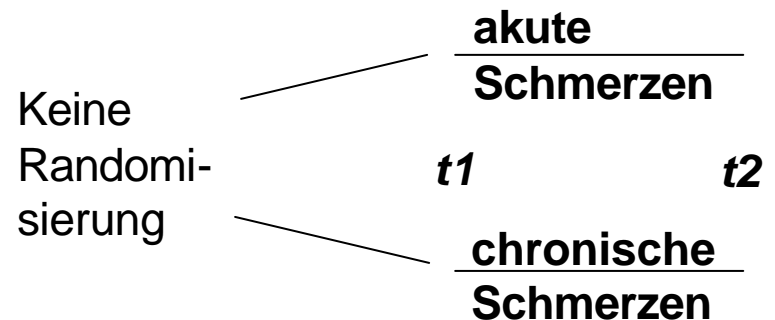
Änderungssensitivität – Messdesigns (2)

Zwei-Gruppen-Designs zur Ermittlung der Änderungssensitivität

- mit Interventions- u. Kontrollgruppe



- zwei wirksame Interventionen mit bekanntem unterschiedlichen Verlauf



Änderungssensitivität – Kennwerte (1)

- **Inferenzstatische Kennwerte:** t-Werte, F-Werte
 - Vorteil: einfach anwendbar
 - Nachteil: abhängig von Stichprobengröße
- **Korrelationen** mit veränderlichen Außenkriterien:
 - Vorteil: Berechnungen zu Sensitivität und Spezifität möglich
 - Nachteil: Schwierigkeit einen validen „Gold-Standard“ zu finden z.B. Arzttrating oder Patientenrating

Änderungssensitivität – Kennwerte (2)

- **Effektstärken (mehrere Varianten):**
 - Vorteile:
 - nicht abhängig von Stichprobengröße
 - Interpretierbarkeit (vgl. Effektstärke von Cohen)
 - Nachteile:
 - unterschiedliche Ergebnisse möglich abhängig von der Effektstärke

Änderungssensitivität – Kennwerte (3)

- **Prozentanteil von Personen mit reliabler Veränderung**
 - **Vorteile:**
 - ermöglicht relativen Vergleich von Fragebögen
 - berücksichtigt unterschiedliche Reliabilität der Fragebögen
 - **Nachteile:**
 - Kennwerte nur relativ interpretierbar
 - Voraussetzung: Reliabilität $r(xx)$ muss bekannt sein, Unkorreliertheit der Messfehler, $sd(t1) = sd(t2)$

Effektstärken der Änderungssensitivität

- Prinzip der Effektstärken:
 - „Standardisierung“ der Mittelwertdifferenz $M(t_2) - M(t_1)$ an der Streuung $sd(x)$
 - Verwendung unterschiedlicher Streuungen $sd(x)$

$$ES = \frac{M_{t_2} - M_{t_1}}{sd_x}$$

Effektstärken der Änderungssensitivität

$$SES = \frac{M_{t2} - M_{t1}}{sd_{x(t1)}}$$

$$SRM = \frac{M_{t2} - M_{t1}}{sd_{d12}}$$

$$GRI = \frac{M_{t2} - M_{t1}}{sd_{d01}}$$

- **SES** = *standardized effect size*
- Standardabweichung der Messwerte $x(t1)$
- hohe Werte bei kleiner Ausgangsvariabilität

- **SRM** = *standardized response mean*
- Standardabweichung der Messwerte $x(t2) - x(t1)$
- hohe Werte bei kleiner Variabilität der Differenzen $x(t2) - x(t1)$

- **GRI** = *Guyatt's responsiveness index*
- Standardabweichung der Messwerte $x(t1) - x(t0)$
- hohe Werte bei kleiner Variabilität der Differenzen $x(t1) - x(t0)$

Zusammenfassung

- Klassische Gütekriterien bzw. Kennwerte nicht ausreichend
- Auswahl alternativer Designs und Kennwerte für die Erfassung der Änderungssensitivität vorhanden
- Mehrere Kennwerte der Änderungssensitivität berichten

*Vielen Dank für Ihre
Aufmerksamkeit!*

Kontakt: wilmar.igl@mail.uni-wuerzburg.de

Anhang: Literatur

1. Guyatt, G. H., Kirshner, B. & Jaeschke, R. (1992). Measuring health status: What are the necessary measurement properties? *Journal of Clinical Epidemiology*, 45, 1341–1345.
2. Guyatt, G., Walter, S. & Norman, G. (1987). Measuring change over time: Assessing the use-fulness of evaluative instruments. *Journal of Chronic Disease*, 40, 171–178.
3. Schuck, P. (2000). Designs und Kennziffern zur Ermittlung der Änderungssensitivität von Fragebogen in der gesundheitsbezogenen Lebensqualitätsforschung. *Zeitschrift für Medizinische Psychologie*, 9, 125–130.

Anhang: Reliable Veränderungswerte

$$\text{SEM} = \text{SD}_x \cdot \sqrt{1 - r_{xx}}$$

SEM ist eine Funktion der Reliabilität(r_{xx}) und der Standardabweichung (SD_x)

$$\text{SRD} = 1,96 \cdot \sqrt{2} \cdot \text{SEM}$$

SRD entspricht dem 95%-Konfidenzintervall des Standardmessfehlers (SEM) der Differenzen

→ **Prozentanteil** von Personen mit reliabler Veränderung

- in positiver Richtung („gebessert“): $\frac{n(d_{12} > \text{SRD})}{N}$
- in negativer Richtung („verschlechtert“): $\frac{n(d_{12} < -\text{SRD})}{N}$