

# **Relevanz von Effektstärken**

**Jürgen Windeler, Stefanie Thomas**

**Medizinischer Dienst des Spitzenverbandes Bund der Krankenkassen (MDS)**

**Berlin, 26.01.2010**

**ganz aktuell**

Januar-Ausgabe 2010

„... war weder das absolute Risiko für nicht tödliche Herzinfarkte noch für die koronare Herzkrankheit **klinisch relevant** vermindert – obwohl die Unterschiede statistisch signifikant sind [...]

Eine NNT=200 bedeutet doch, dass 200 Patienten über viele Jahre einer intensiven Therapie unterzogen wurden, aber im Vergleich zur Standardtherapie nur bei einem Patienten ein tödlicher Herzinfarkt verhindert werden konnte.“



## **DER SPIEGEL**

### **Pillen zum Vergessen**

Neuartige Medikamente werden als Hoffnung gegen Alzheimer gepriesen. Doch neue Studien kommen zu dem Befund: Die teuren Pillen helfen den Patienten nicht.

## aus MDS-Gutachten 2001

zu Cholinesterase-Hemmern:

„Zur Beurteilung ihrer klinischen Relevanz wurde eine Experten-Runde eingerichtet. Die Experten beurteilen die positiven Effekte der 3 Antidementiva **als klinisch relevant**, wenn auch das Ausmaß **im Mittel als gering** angesehen wird. **In Einzelfällen** können die positiven Änderungen nach Aussagen der Experten sehr ausgeprägt sein, dem steht ein **Anteil von Nonrespondern von ca. 30%** gegenüber.“

## aus IQWiG-Bericht 2009

zu Memantine:

„ .... Für die Einschätzung der **Relevanz eines Gruppenunterschieds** bezüglich stetiger Skalen sollten aus der Literatur bekannte „Minimal Important Differences“ (MID) für den Gruppenunterschied berücksichtigt werden. Da für keine der eingesetzten Skalen eine verwendbare MID für den Gruppenunterschied vorlag, wurde eine Berechnung von Cohen's d durchgeführt. Für Cohen's d wurde eine Relevanzgrenze von 0,2 angenommen, ...“

## These

- die Vielschichtigkeit des Begriffes (klinische) Relevanz
- die Ungenauigkeit von Definitionen und Konzepten
- die Vielzahl von Lösungsvorschlägen,
- die kaum praktisch eingesetzt werden

legen nahe,

dass es für die Bewertung der Relevanz keine triviale Lösung gibt,

oder jedenfalls, dass bisher Herangehensweisen und Lösungsvorschläge unbefriedigend sind

## Äquivalenz (Nicht-Unterlegenheit)

### Continuation of Low-Dose Aspirin Therapy in Peptic Ulcer Bleeding

#### A Randomized Trial

Joseph J.Y. Sung, MD, PhD; James Y.W. Lau, MD; Jessica Y.L. Ching, MPH; Justin C.Y. Wu, MD; Yuk T. Lee, MD; Philip W.Y. Chiu, MD; Vincent K.S. Leung, MD; Vincent W.S. Wong, MD; and Francis K.L. Chan, MD

*Ann Intern Med.* 2010;152:1-9.

Blutungsrate nach Absetzen von ASS: erwartet 6.7 %

**Irrelevanzbereich: bis 10 %-Punkte Differenz**

Ergebnis: Differenz 4.9 %-Punkte [-3.6 ; 13.4]

### **Festlegungen** (teilweise informell, implizit, pragmatisch)

- Test (Quantifizierung notwendig)
- Grundlagen (Eckpunkte) durch Standardbehandlung
- Diskussion über Abwägungsnotwendigkeit (anderer Vorteil) wird breit geführt


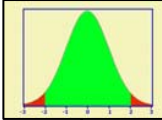


## Fragen

### Suche nach der „minimalen relevanten Differenz“: MRD

- ⇒ **Gibt es Methoden**, a priori eine Grenze für klinisch relevante Therapieeffekte festzulegen?
- ⇒ Falls ja - **Wie funktionieren** solche Methoden?
- ⇒ **Wie valide** sind sie?
- ⇒ Wie sind resultierende MRDs **bei der Relevanzbewertung /-prüfung methodisch korrekt anzuwenden?**




# Systematisierung

Methodentyp	
<b>PbM</b>	 Präzisionsbasierte Methoden
<b>VbM</b>	 Verteilungsbasierte Methoden
<b>AbM</b>	 Ankerbasierte Methoden
<b>MbM</b>	 Meinungsbasierte Methoden

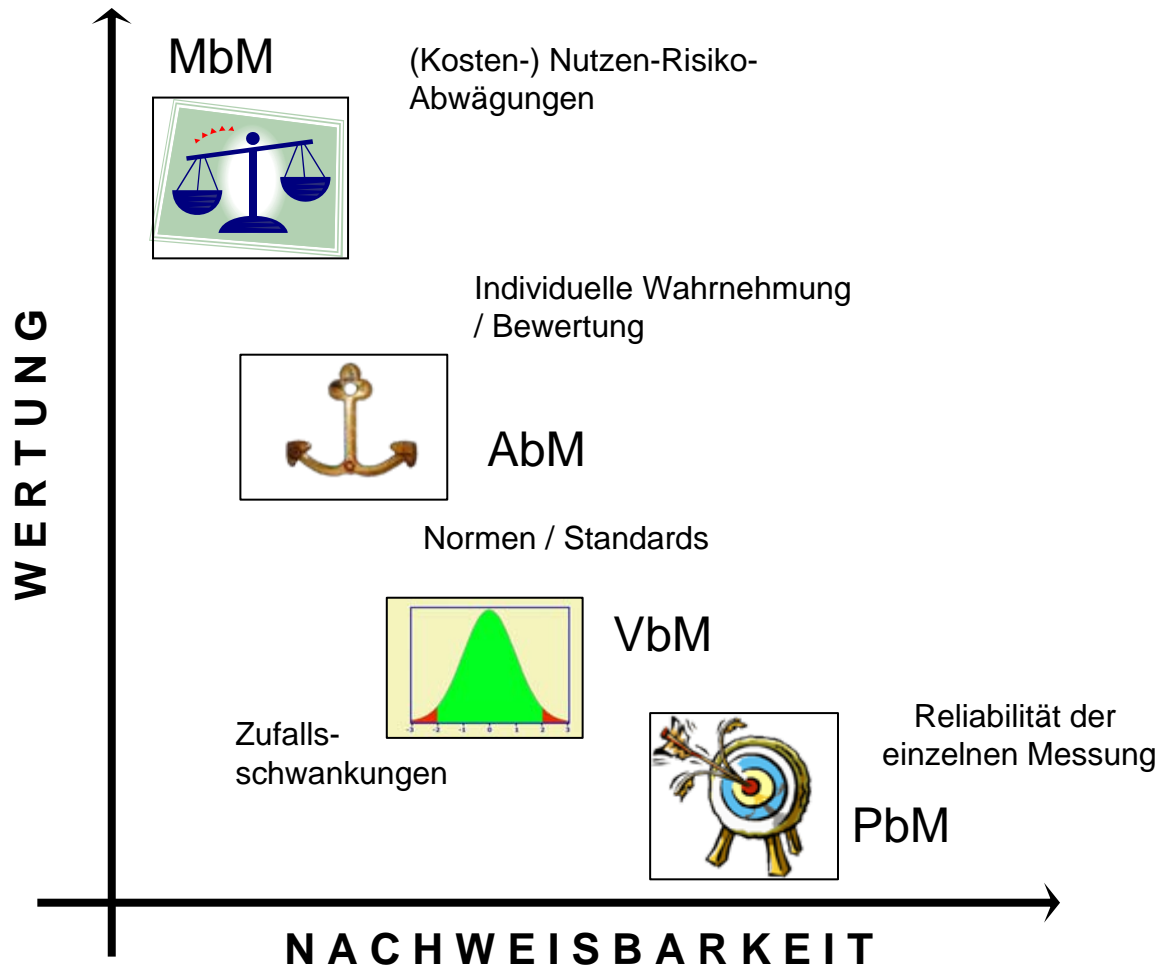
Thomas, S. Klinische Relevanz von Therapieeffekten - Systematische Sichtung, Systematisierung und Bewertung methodischer Konzepte. Dissertation, Universität Essen, 2009

# Systematisierung

	Methoden-Prinzip
	Berücksichtigung der Mess(un)genauigkeit aktueller Veränderungsmessungen
	„Normen“ Abweichung vom Populationen in Richtung
	Verankerung des Relevanz aussage-kräftig „Konstrukt-Validität“
	Festlegung durch (Experten-) Befragung

**Typ-spezifische „Stellschrauben“  
 mit definierten  
 Variationsmöglichkeiten**

# Systematisierung



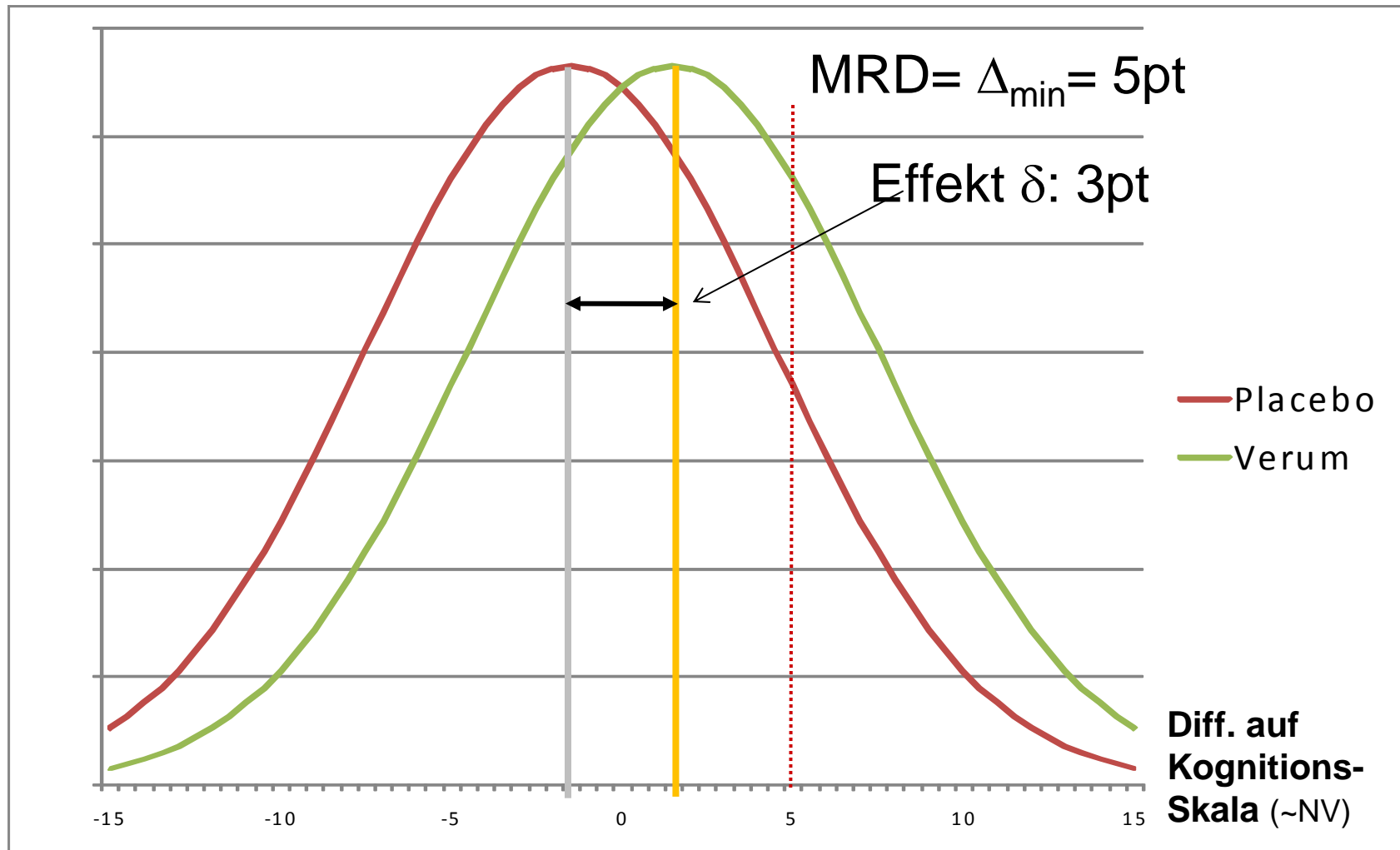
## AD2000 (2004)

### Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial

Do the significant but small improvements in cognition and functionality seen over the 2-year study (average 0.8 MMSE points and 1.0 BADLS points) lead to a worthwhile improvement in health-related quality of life? Definition of a worthwhile clinical response, or the minimum clinically important difference, on simple composite outcome scales is difficult.<sup>34</sup> For example, the US Food and Drug Administration have suggested that a change equates to 1.4 MMSE points since the mean annual rate of decline in Alzheimer's disease is reported to be about 2.8 points (range 1.8–4.2).<sup>36</sup> However, a large survey of clinicians revealed that 45% judged a 3-point change as the minimum clinical benchmark.<sup>36</sup> The upper confidence limit for the improvement in MMSE from AD2000 is 1.2 points, which is less than both of these estimates. The benefits of 10 mg donepezil seem slightly larger than the average benefits for 5 mg and 10 mg reported here, but not enough to reach the minimum clinically important difference threshold.

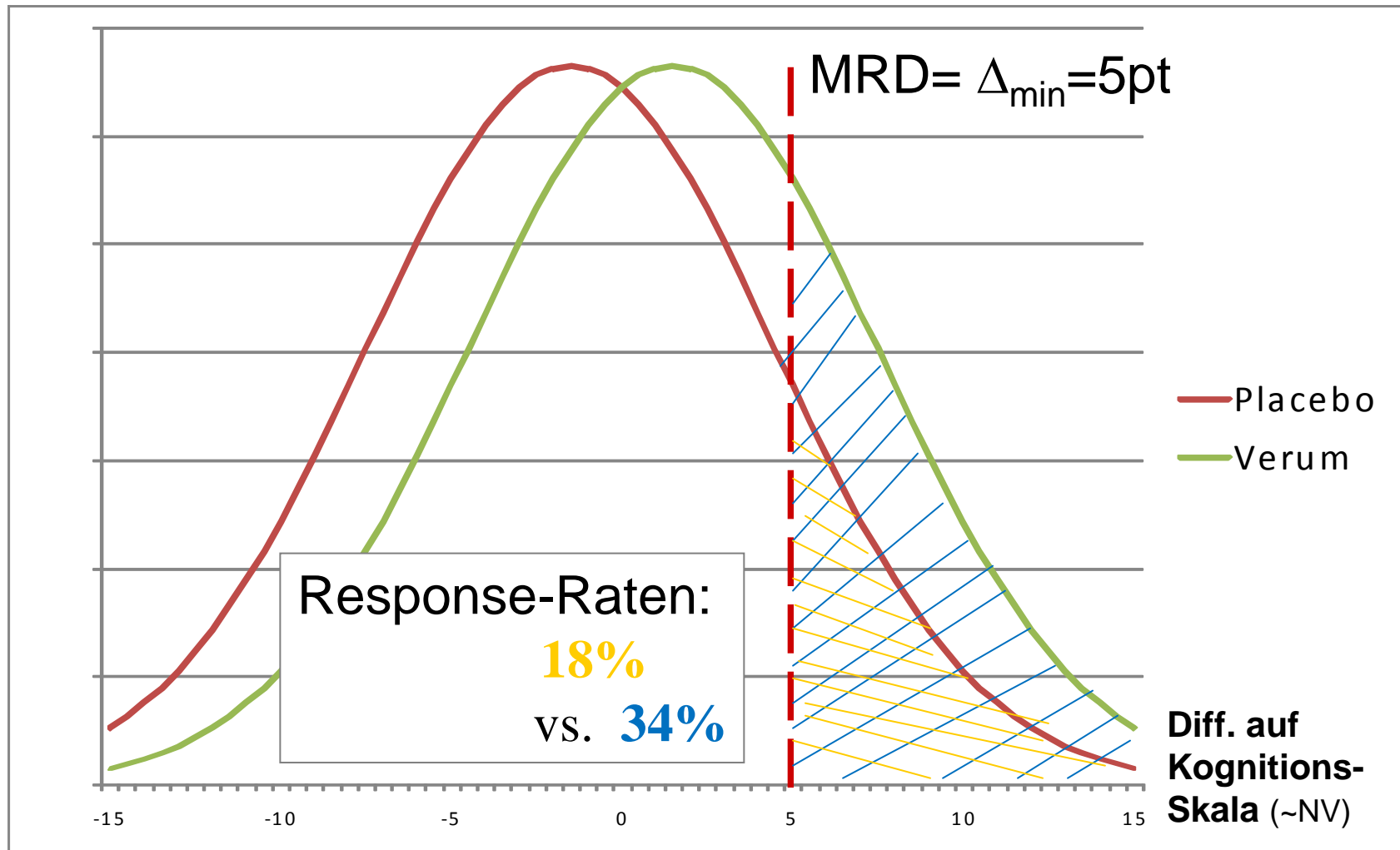
## Einzel vs. Gruppe

Fälschliche Interpretation eines Effekts mittels individueller MRD



## Einzel vs. Gruppe

Fälschliche Interpretation eines Effekts mittels individueller MRD



## Einzel vs. Gruppe

Es gibt Vorschläge, wie ein individueller relevanter Response definiert werden kann

**aber**

keine Hinweise, wie von der individuell relevanten Veränderung auf die Effektbewertung übergegangen werden soll

**Warum ist das eine offenbar einfacher als das andere ?**

**These** (durch Sichtung methodischer Konzepte gestützt):

**Die Grundlage jeder Relevanzdiskussion ist ein Werturteil.**

Auf dessen Basis kann das Problem formal behandelt werden.

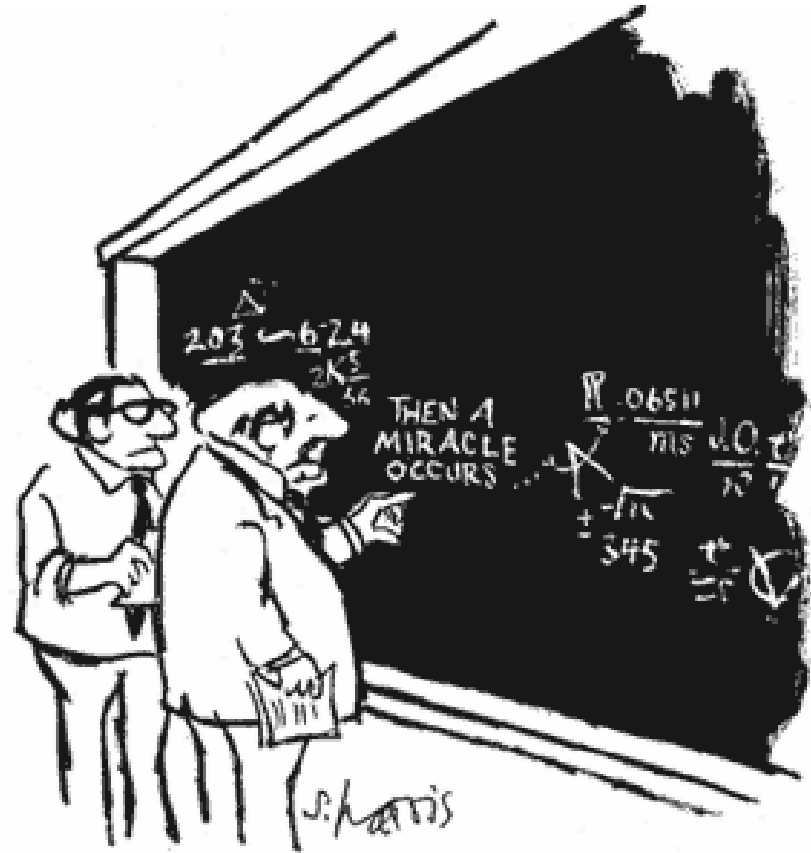
Individuell kann dies noch gelöst werden (oder man tut so, als ob), auf der abstrakten Gruppenebene nicht (so leicht).

## Einzel vs. Gruppe

Die Schwierigkeit der Lösung für

- formale Lösungsansätze erdrücken das Grundproblem nicht berücksichtigt
- das Problem ignoriert wird
- die Diskussion durch wiederholte, ungeprägt ist, die u.a. die Aspekte individuelle vs. Gruppen Effektschätzer Fallzahlplanung

umfassen



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."



## Fazit

„Further research is needed“ ... ☹️

1. Multidimensionalität von Relevanz-Bewertungen in die **Diskussion** holen
  - Fragen zur Zahlungsbereitschaft von der klinischen Relevanzbewertung trennen
2. **Forschungsbedarf**
  1. Verfahren zur Etablierung von Responsekriterien bei Gesundheitsskalen
  2. Methoden bei Betrachtung multipler Endpunkte
  3. Einbezug von Entscheidern / Patienten

**aber:** Institutionen im deutschen Gesundheitswesen sind **heute** gefordert, Relevanzbewertungen abzugeben ...

## Optionen

1. (allgemeine) quantitative Regel, z.B. StDiff 0.x; RRR 0.x
2. enger (formal orientierter) Rahmen
  - a. Betrachtung mehrerer Endpunkte
  - b. Einbezug von Entscheidern / Patienten
3. weiter (kriterien-basierter) Rahmen (z.B. Schwere der Erkrankung, Behandlungsalternativen, Sicherheit)
4. situationsbezogene Abwägung / Bewertung

**basierend auf einem Werturteil !**