

# **Methodische Aspekte bei der Analyse von Daten zur Lebensqualität in Nutzenbewertungen**

**Christoph Schürmann**

# Zentrale Fragen

**Was und womit messen ?**

(Endpunkt, Messinstrument)

**Wann und wie oft messen ?**

(Zeitpunkte)

**Wie auswerten ?**

(Operationalisierung, Effektmass, Effektschätzung)

**Wie bewerten ?**

(Beurteilung Effektgröße, Ergebnissicherheit)

I.

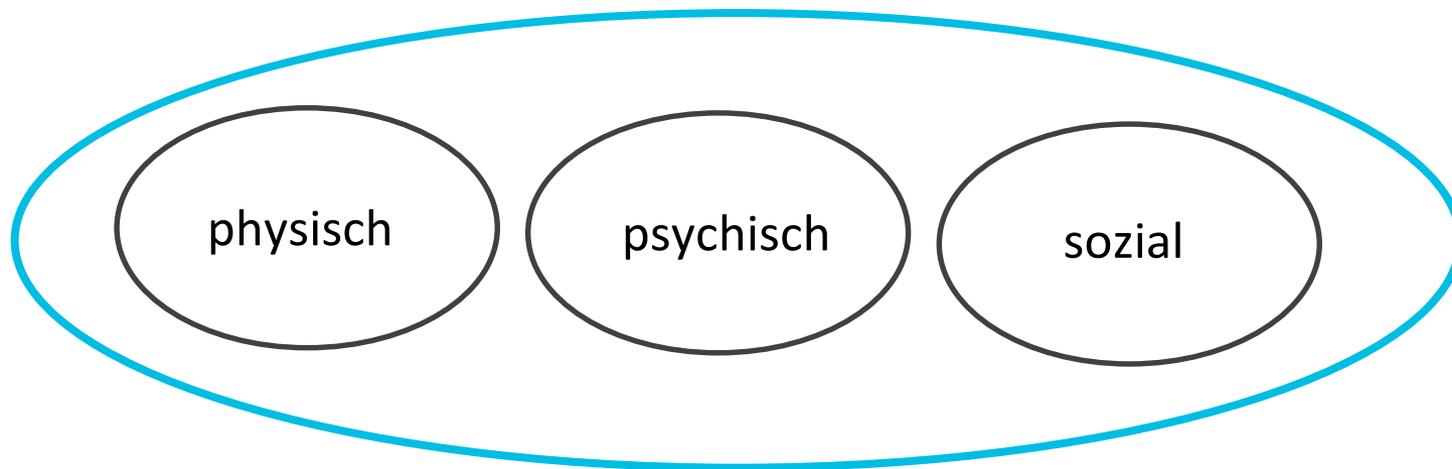
# Was und womit messen?

# Gesundheitsbezogene Lebensqualität

- Konstrukt aus mehreren Dimensionen
- Schwer objektivierbar
- Subjektivität des Patientenerleben
- Messung direkt durch Betroffene (patient reported outcome)
- Skalenbasierte Instrumente notwendig

# Komponenten der Lebensqualität

HRQoL



[WHO, 1995]

# Verwendbare Skalen zur HRQoL

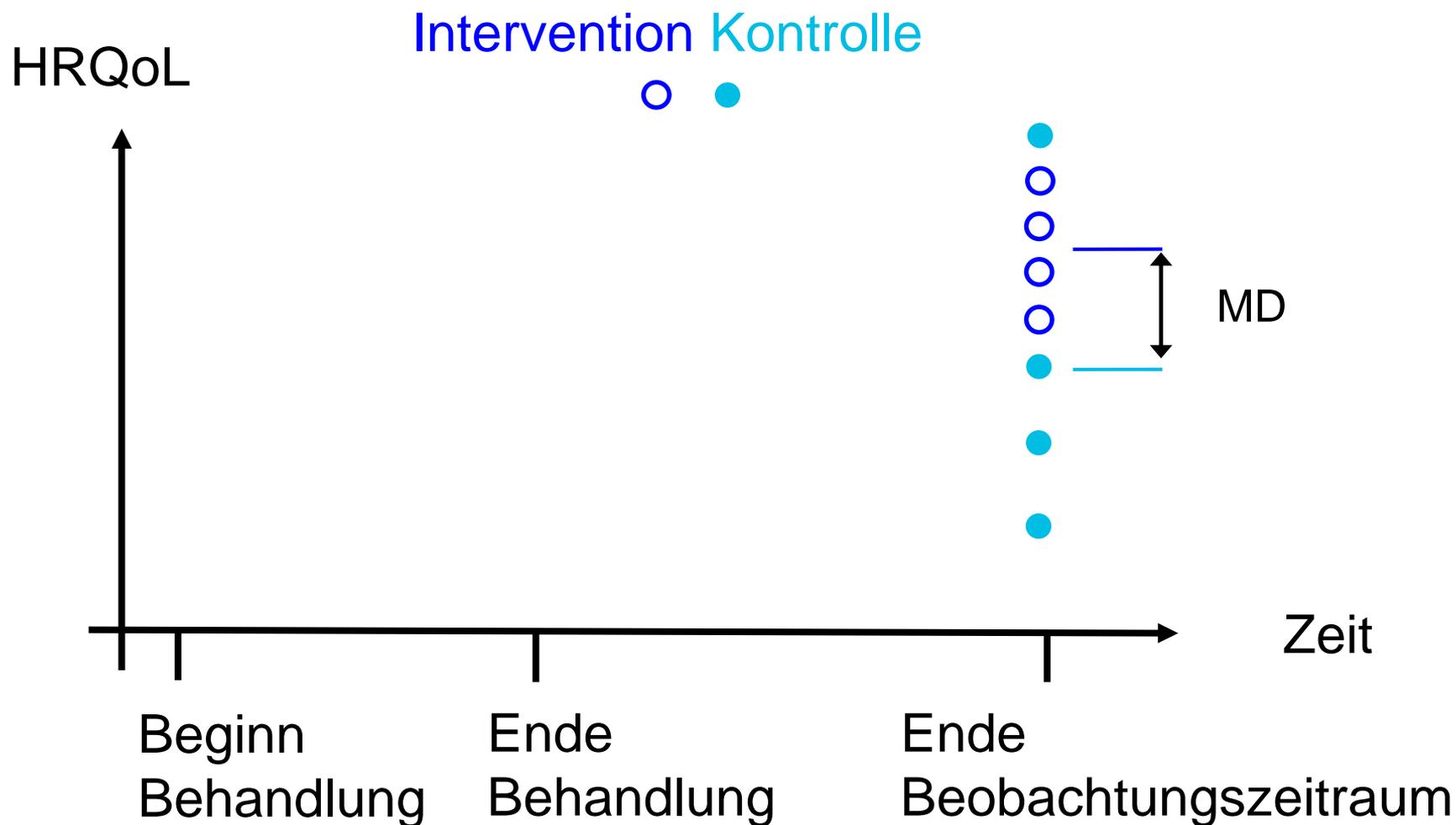
- Berücksichtigung aller Dimensionen
- Allgemein (indikationsübergreifend) oder spezifisch
- Validiert oder etabliert
  
- Beispiele allgemein: SF-36
- Beispiele spezifisch: EORTC QLQ-C30, FACT-G

## II.

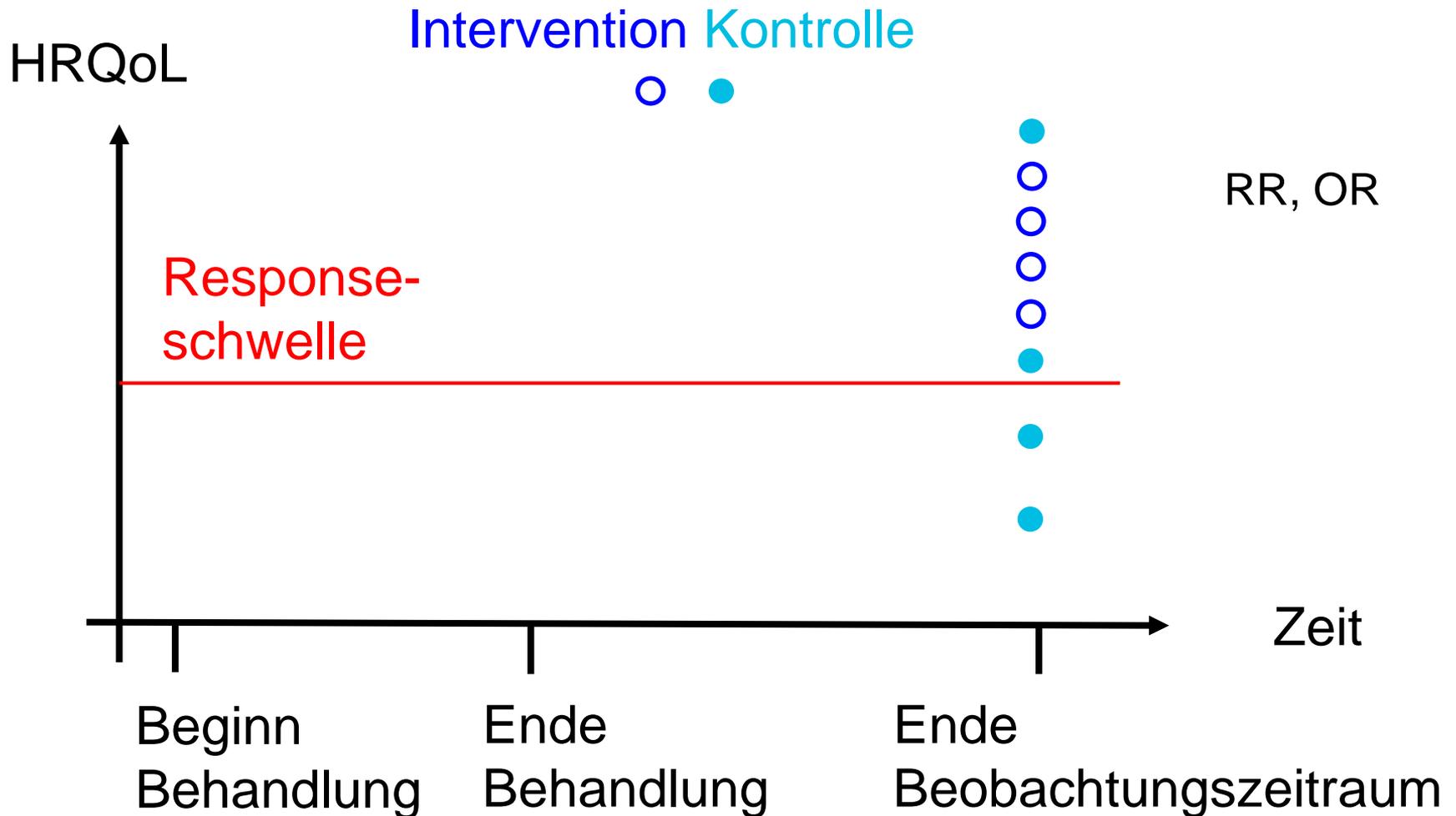
Wann und wie oft messen?

Wie auswerten?

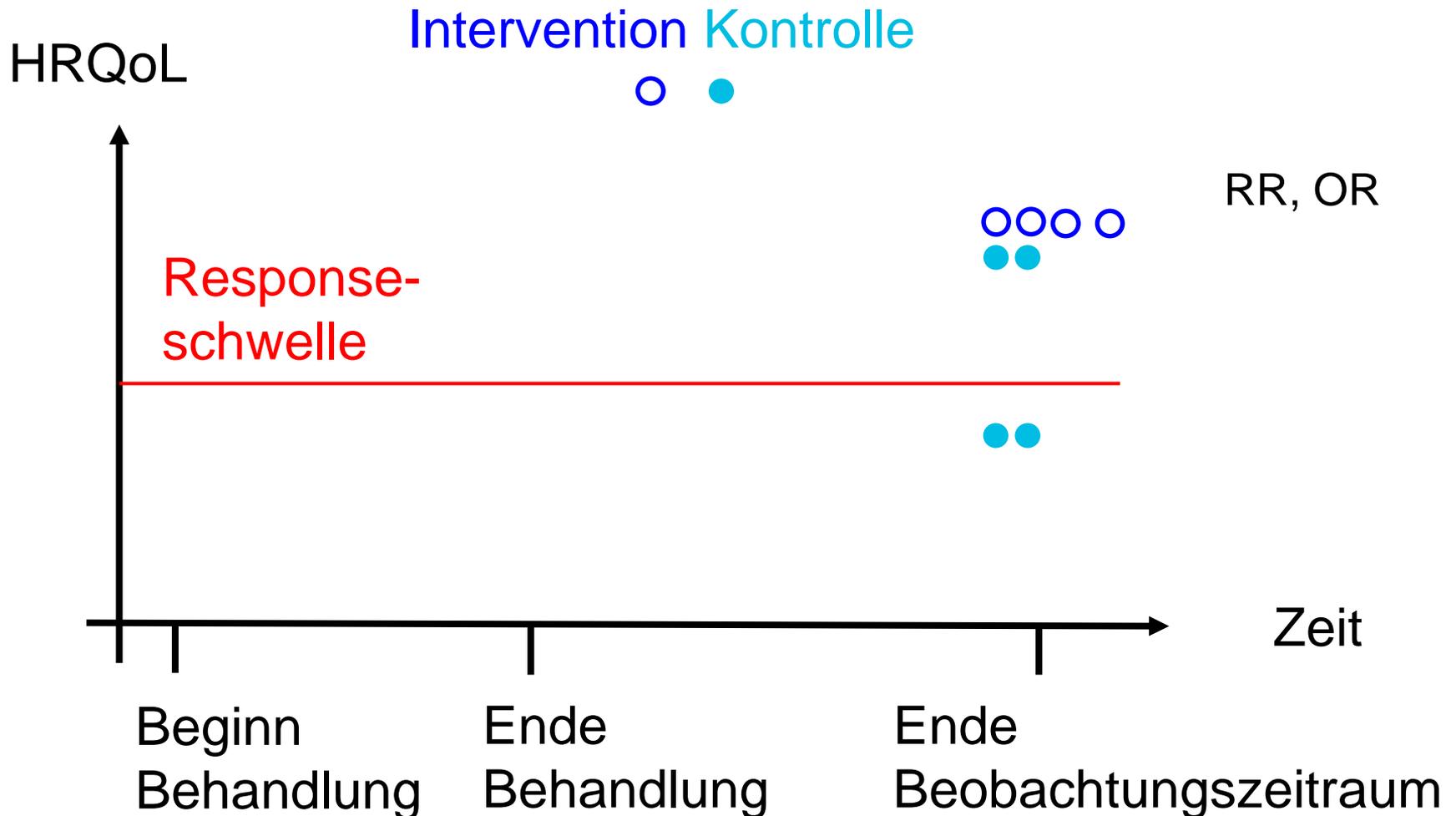
# Einmalige Messung



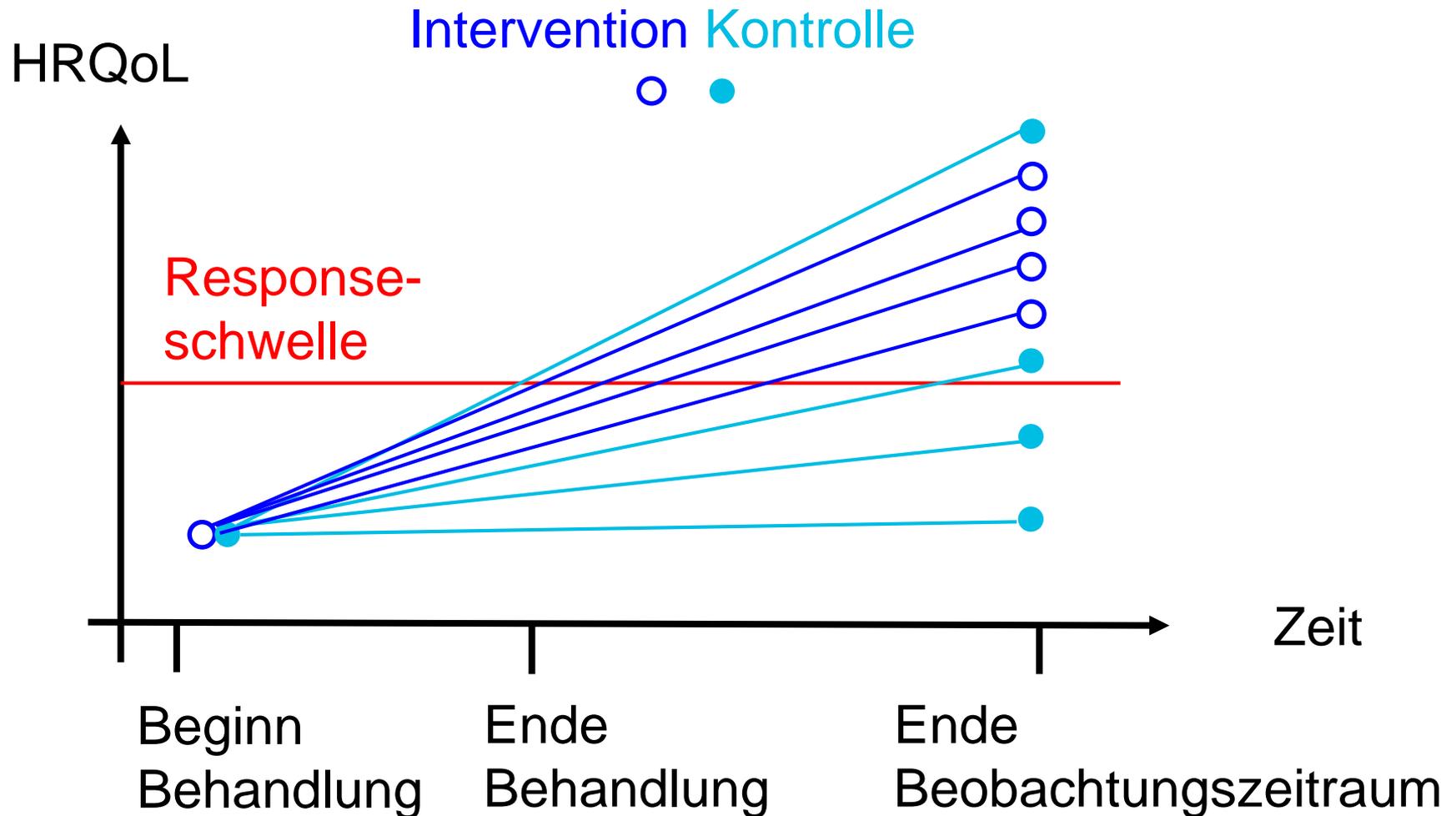
# Einmalige Messung



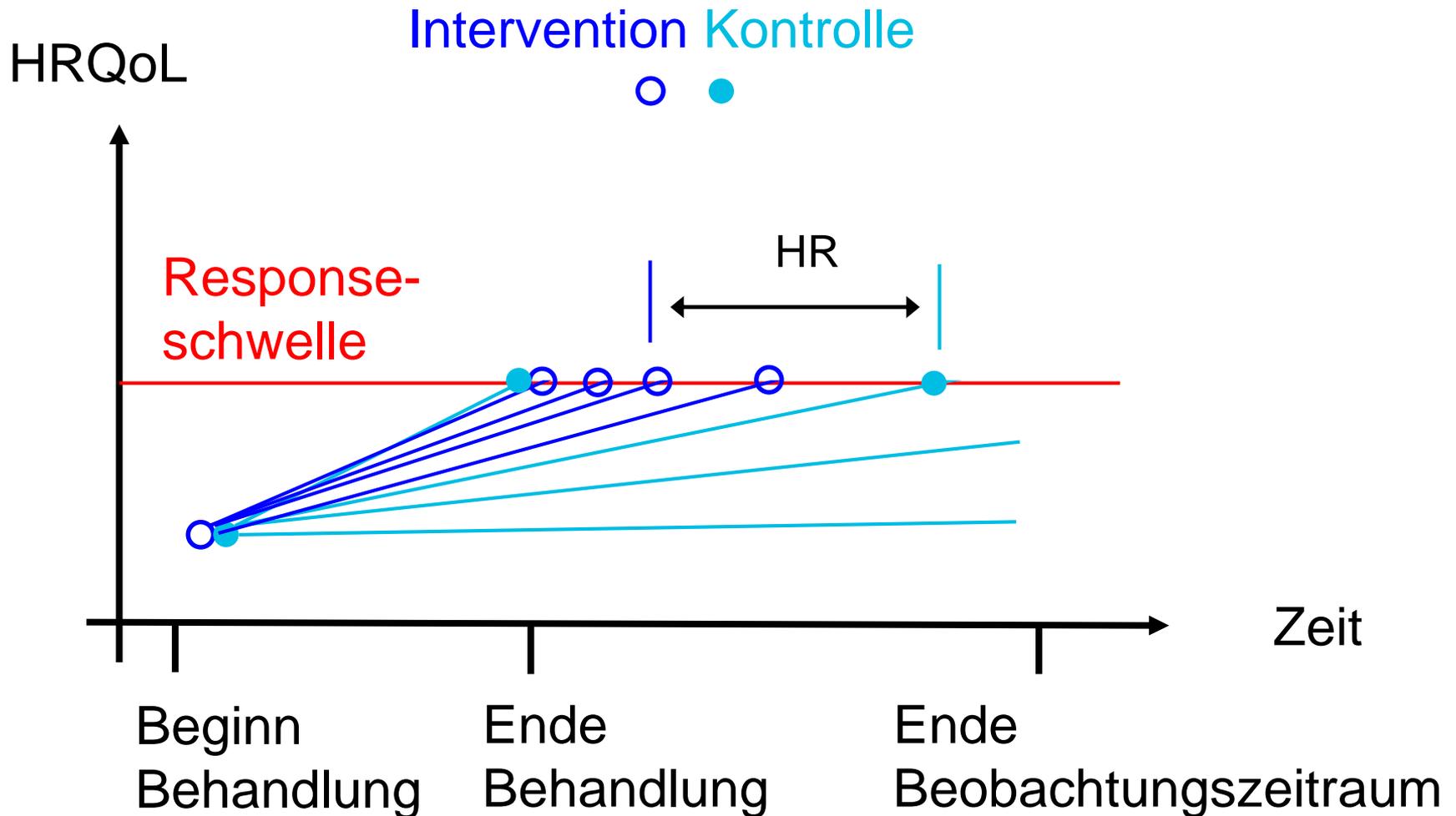
# Einmalige Messung, dichotom



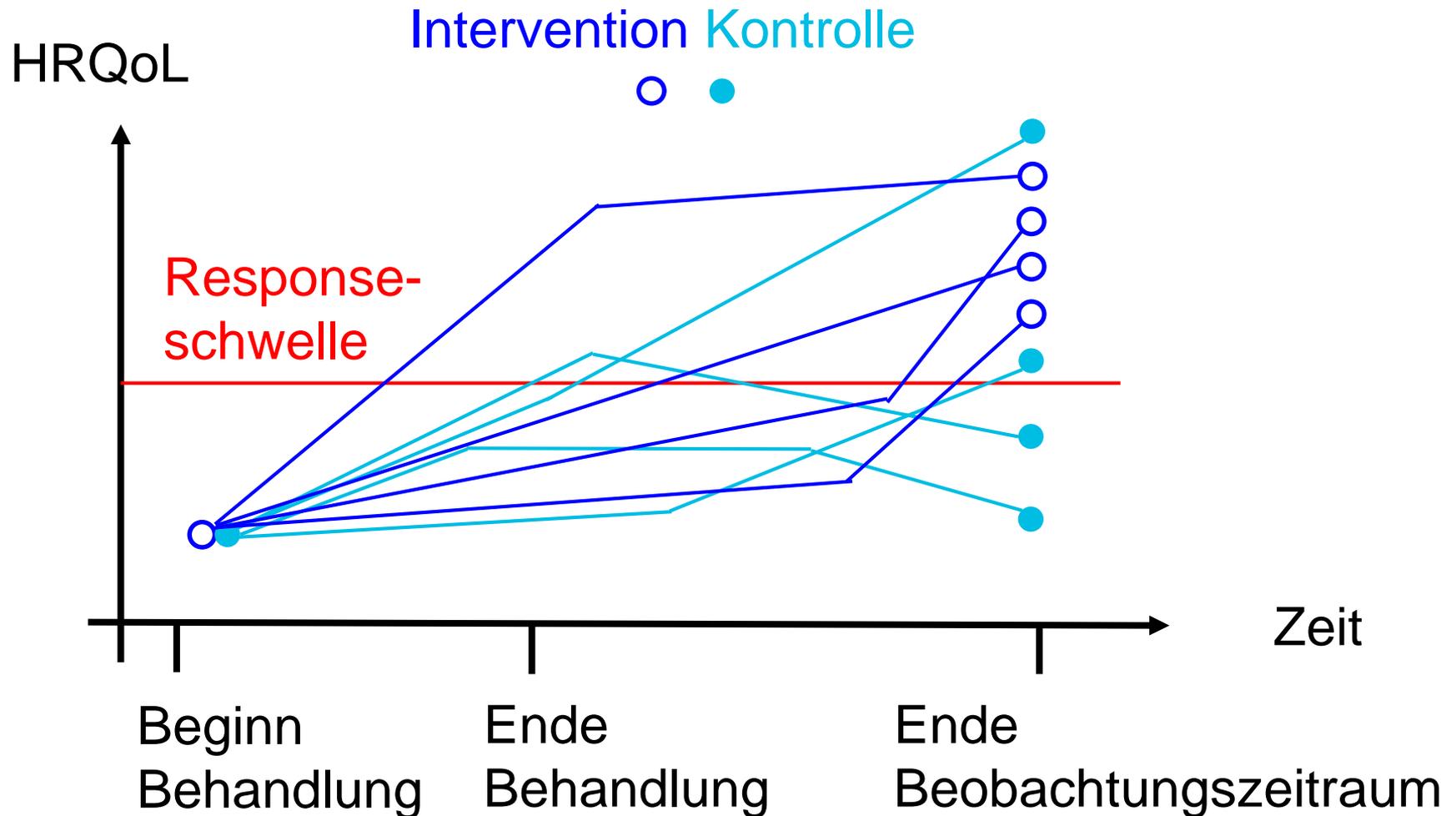
# Monotone Verläufe



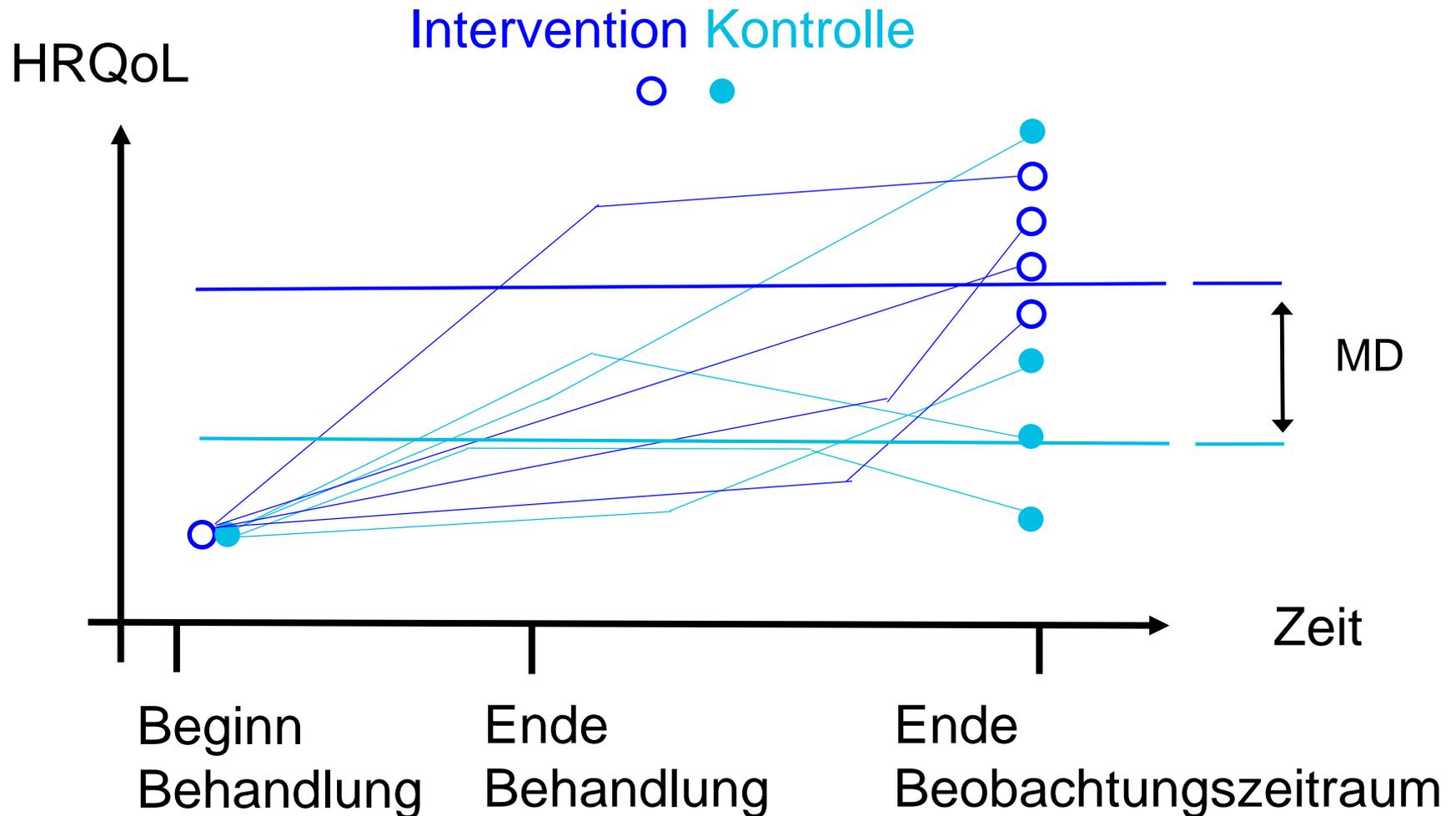
# Monotone Verläufe



# Nicht monotone Verläufe



# Mehrfache Messung



# III.

Wie auswerten?

# Ziel der Bewertung

- Den gesamten Beobachtungszeitraum abdecken  
(gleichlang in beiden Gruppen, unabhängig von Zwischenereignissen)
- Die gesamte Information des Zeitraums berücksichtigen
- Estimand: Treatment Policy
- Auswertung des Gesamtzeitraums  
oder
- Auswertungen zu mehreren Zeitpunkten/Abschnitten

# Stetige Operationalisierung

- Mittelwerte oder Summe der Skalenwerte
- Modelle mit Meßwiederholung (repeated measurements)
- Ggf. kumuliert über die Zeit - area under the curve (AUC)
- Effektmaß: Mittelwertdifferenz, standardisiert (Hedges'  $g$ )
- Interpretierbar mittels Irrelevanzschwelle

# Binäre Operationalisierung

- Einteilung der Skalenwerte mittels Responseschwelle
- Auswertung als Ereigniszeitanalyse (HR) oder zu mehreren Zeitpunkten (RR)
- Interpretierbarkeit der Responseschwelle:  
Bedeutung für die Patientinnen?
- Konzepte: MID, MCID
- MID spezifisch für Skalen und Anwendungsgebiete

# Was ist die "richtige" MID?

- Eigene Beobachtung [IQWiG, 2020a]:
  - hohe Variabilität bei MIDs, fehlende Standards
  - meist unterhalb 20% Skalenspannweite
- Literatur [Carrasco-Labra, 2021]:
  - schwache / unklare Korrelation mit Ankern
  - hohe Variabilität bei MIDs
  - fehlende Angaben zur Präzision
  - verschiedene methodische Ansätze
- Standardisierte methodische Bewertung [Devji, 2020]
  - in der Praxis nicht umsetzbar

# Responseschwelle – neuer Umgang

- Ziele: Kleine Veränderungen abbilden  
Risiko für ergebnisgesteuertes Berichten minimieren
- Regelmäßiger Schwellenwert statt MID [IQWiG, 2020b]:
  - > 15% der Skalenspannweite (vorab festgelegt)
  - = 15% der Skalenspannweite (post hoc)
- 15% als spürbare Veränderung hinreichend sicher
- nicht für Skalen mit natürlichen Einheiten
- nicht für unbegrenzte Skalen

## Besondere Fälle: SF-36

- generisches Instrument, 8 Domänen
- 2 Summenscores: PCS, MCS
- Theoretisch konstruierbare Extremwerte:  
PCS: 5 – 80    MCS: -3 – 80
- Extremwerte der Normstichprobe 2009:  
PCS: 7 – 70    MCS: 6 – 70
- 15% bezüglich realisierter Spannweite der Normstichprobe  
=> Responseschwelle: 10 Punkte

## Besondere Fälle: EORTC

- QLQ-C30 und ggf. Zusatzmodule
- 5 Funktionsskalen zur HRQoL
- Spannweite 100, aber diskrete Ausprägungen
- Etabliert: MID 10 (10% der Spannweite) [Osoba, 1998]
  
- Kriterium 15% führt oft zu gleichen Häufigkeiten  
=>10 Punkte sind hinreichende Annäherung
  
- Schwellenwert 10 Punkte auch bei Zusatzmodulen,  
falls mehrheitlich (> 70%) gleiche Häufigkeiten resultieren

IV.

Wie bewerten?

# Verzerrungsaspekte

- Studiendesign:
  - Fehlende Verblindung
  - nicht zulässige Behandlungswechsel
  - Unterschiedliche Beobachtungsdauern (zwischen oder innerhalb von Armen)
- Studienverlauf: Patientenbedingt fehlende Werte
- Auswertung:
  - Ergebnigesteuerte Berichterstattung
  - Ersetzungsverfahren
  - Sensitivitätsanalysen

# Anforderungen an die Berichterstattung

- Darstellung der Rückläufe mit Angaben zu auswertbaren Fragebögen und Patientinnen unter Beobachtung
- Darstellung von Verlaufskurven / Zwischenauswertungen über die Zeit
- Angaben zur Beobachtungszeit
- Gründe für fehlende Daten
- Angaben zur geplanten Auswertung
- Darstellung von Gesamtscore und Subskalen

V.

# Beispiele

## Atezolizumab (A20-97)

- Hepatozelluläres Karzinom (fortgeschritten, nicht resezierbar)
- Atezolizumab + Bevacizumab vs. Sorafenib
  
- 1 Studie, nicht verblindet
- Beobachtungsdauer: ~ 8,5 Monate [Median]
- Patienten: 336 vs. 165
- Wiederholte LQ-Erhebung zum Zyklusende  
bis 1 Jahr nach Abbruch Studienmedikation oder Progression
  
- EORTC QLQ-C30 und QLQ-HCC18

## Atezolizumab – EORTC QLQ

- Responseschwelle: Verschlechterung um 10 Punkte  
(hinreichende Annäherung an 15% Skalenspannweite)
- Verfügbare Auswertungen: Überlebenszeitanalysen
  - für Verschlechterung
  - für bestätigte Verschlechterung
- Hohes Verzerrungspotenzial

# Atezolizumab: Bewertung

- Funktionsskalen des QLQ-C30 (Auswahl):
  - emotionale Funktion: HR = 0,45 [0,34; 0,59]  
=> Anhaltspunkt: erheblich
  - kognitive Funktion: HR = 0,66 [0,52; 0,85]  
=> Anhaltspunkt: beträchtlich
  
- Funktionsskalen des QLQ-HCC18 (Auswahl):
  - Ernährung: HR = 0,49 [0,39; 0,62]  
=> Anhaltspunkt: erheblich
  - Sexualeben: HR = 0,85 [0,63; 1,15]  
=> kein Anhaltspunkt

## Risankizumab (A19-41, A19-87)

- Plaque-Psoriasis
- Erstmalige systemische, keine konventionelle Therapie
- Risankizumab vs. Ustekinumab
  
- 2 Studien, verblindet
- Dauer: je 52 Wochen
- Patienten: 100 vs. 33 und 90 vs. 36
- Erhebung der LQ bis zum Studienende, wiederholte Messung zu Wochen 0, 12, 16, 52

# Risankizumab - DLQI

- Dermatology Life Quality Index
- Responseschwelle: Erreichen von Score 0 oder 1 (Freiheit von Beeinträchtigungen)
- Verfügbare Auswertungen:
  - Überlebenszeitanalyse für Responder
  - Responderanalyse zu Studienende
  - Mittelwertdifferenz nach Studienende
  - (Addendum) Sensitivitätsanalysen

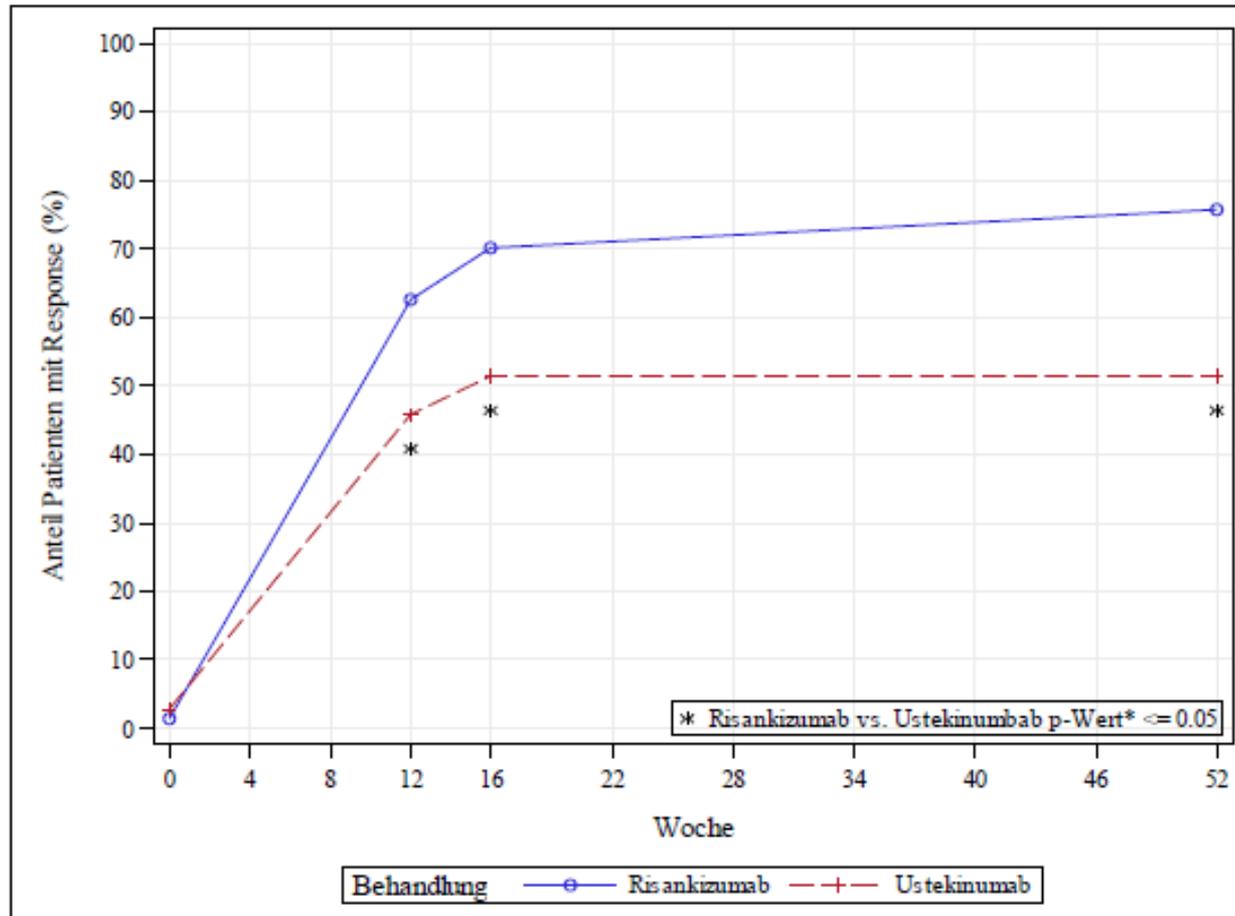
## Risankizumab: DLQI

- Fehlende Beobachtungen zu Studienende  
5% vs. 12% (Studie 1)  
4% vs. 17% (Studie 2)  
=> Relevante Unterschiede zwischen Armen
- Ersetzungsverfahren NRI (non-responder imputation)
- Darstellung von Verlaufskurven

Addendum:

- Sensitivitätsanalysen mit anderen Ersetzungsverfahren

# Risankizumab: Verlaufskurven



# Risankizumab: Bewertung

- NRI: RR = 1,47 [1,16; 1,86]
- Hohes Verzerrungspotenzial  
=> Hinweis; beträchtlich
  
- Addendum:
  - Sens-1: LOCF: RR = 1,38 [1,11; 1,71]
  - Sens-2: MI: RR = 1,42 [1,12; 1,79]
- Hohe Ergebnissicherheit  
=> Beleg; beträchtlich

VI.

Fazit

# Anforderungen an Studienplanung

- Wahl validierter LQ-Instrumente
- Verblindete Erhebung (ideal)
- Mehrfache Erhebung über den gesamten Studienverlauf
- Festlegung von Auswertungsarten vorab
- ggf. Festlegung von geeigneten Responseschwellen
- Erheben von Abbruch-Gründen

# Anforderungen an die Auswertung

- Auswertung unter Einbeziehung aller erhobenen Daten d. h. Berücksichtigung der Messwiederholungen
- Effektmaße: SMD und / oder HR bzw. RR mit vorab definierten Responseschwellen
- Geeigneter Umgang mit fehlenden Werten (Ersetzungsstrategien o. ä.)
- Sensitivitätsanalysen

## Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)



Im Mediapark 8  
D-50670 Cologne

Fon +49 221 35685-0  
Fax +49 221 35685-1

[info@iqwig.de](mailto:info@iqwig.de)

[www.iqwig.de](http://www.iqwig.de)

[www.gesundheitsinformation.de](http://www.gesundheitsinformation.de)

[www.themencheck-medizin.de](http://www.themencheck-medizin.de)

Twitter: @iqwig and @iqwig\_gi

# Literatur

- WHO (1995): The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med, 41(10), 1403–1409.
- Devji, T, Carrasco-Labra, A, et al. (2020): “Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study“, BMJ 2020;369:m1714.
- Carrasco-Labra, A, Devji, T et al. (2021): "Minimal important difference estimates for patient-reported outcomes: A systematic survey", JoCE 133; 61-71.
- IQWiG (2020): Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden 6.0. IQWiG, Köln.  
<https://www.iqwig.de/ueber-uns/methoden/methodenpapier/>
- IQWiG (2020): Allgemeine Methoden, Version 6.0 vom 05.11.2020. IQWiG, Köln.  
<https://www.iqwig.de/de/methoden/methodenpapier.3020.html>