



# Aussagesicherheit bei der Nutzenbewertung – Praxiserfahrungen

Stefanie Thomas



- Wo „greift“ die Methodik?

## ... Motivation:

- Regeln statt Willkür
- Ableitung eindeutiger Aussagen.

## Was ist neu?

- ...
- Update und Umbenennung des Abschnitts 3.1: Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens
- Löschen des Abschnitts zur Ergebnissicherheit  
Zusammenfassende Bewertung
- ...
- Änderungen und Ergänzungen in Abschnitt 7.3.8: M  
○ Unterteilung in Unterabschnitte  
○ Neuer Unterabschnitt zu Subgruppenanalysen im Rahmen von Meta-Analysen
- ...
- Neuer Abschnitt 7.3.9: Indirekte Vergleiche
- ...

**Surrogate**

**Systematik für Belege,  
Hinweise und Anhaltspunkte**

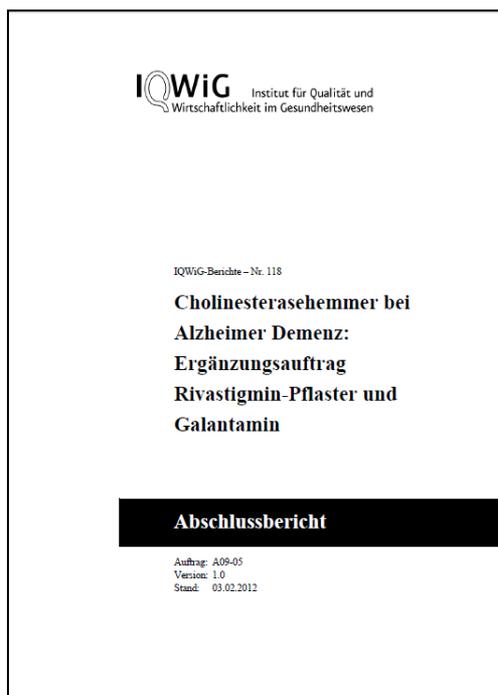
**Konsistenzbetrachtungen  
und Subgruppen**

**Indirekte  
Vergleiche**

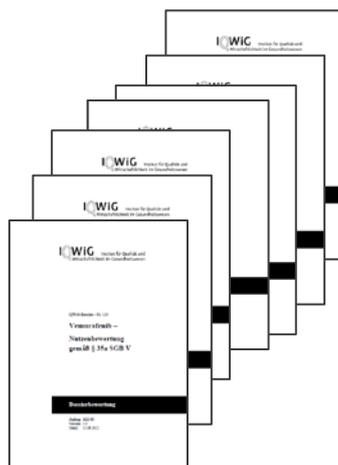
„**Welche** Konkretisierungen sind von Bedeutung?  
Machen sie **Sinn**“ ?

→ Darstellung von Ergebnissen aus der Praxis

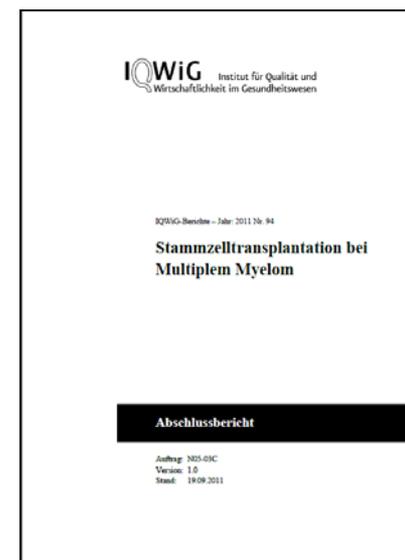
## ➤ Bericht aus Ressort Arzneimittel: Cholinesterasehemmer / Alzheimer



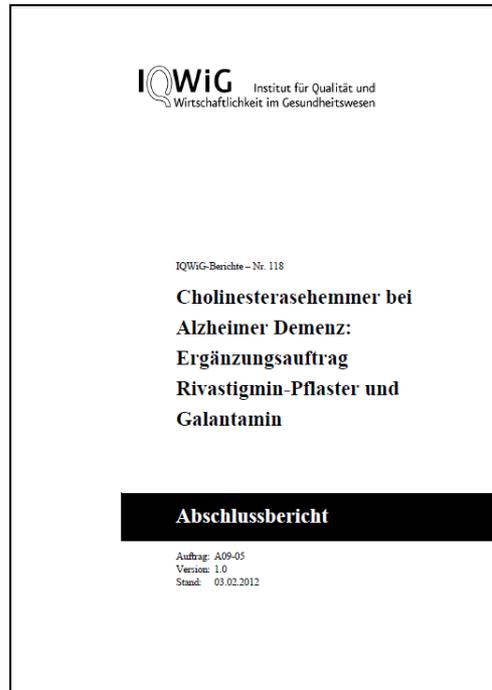
## Ergänzend: Dossierbewertungen



**Abschließend**  
Nichtmedikamentöse  
Therapeutische Intervention:  
Stammzelltransplantation



## ➤ Bericht aus Ressort Arzneimittel: Cholinesterasehemmer / Alzheimer



## Vergleiche

- Galantamin vs. Placebo (the
- Galantamin-Weiterbehandlung vs. Placebo
- Rivastigminpflaster 10cm vs. Placebo

Gemäß „Abschnitt 3.1.4  
Zusammenfassende Bewertung“:  
➤ Welche Nutzensaussagen  
waren (nicht) möglich?  
➤ Warum?

## Wesentliche Endpunkte

- Kognitive Leistungsfähigkeit (**ADAS-cog**)
- Aktivitäten des täglichen Lebens (*versch. Scores*)
- Unerwünschte Ereignisse
- (...)



\* Ergänzungsauftrag!

(1) **Verzerrungspotenzial (VP)** → qualitative Ergebnissicherheit pro Studie/Endpunkt bestimmen,

- **hohe Ergebnissicherheit:** Ergebnis eines RCT mit niedrigem Verzerrungspotenzial.
- **mäßige Ergebnissicherheit:** Ergebnis eines RCT mit hohem Verzerrungspotenzial.
- **geringe Ergebnissicherheit:** Ergebnis einer nicht randomisiert vergleichenden Studie.“

(2) ...**weitere Faktoren (WF)** → ....überprüfen, ggf. modifizieren.

„In begründeten Fällen beeinflussen **weitere Faktoren**  Einschätzungen. Die Betrachtung von **Surrogatendpunkten** (vgl. Abschnitt 3.1.2) oder **Designmängel** bei einer Studie kann z. B. zu einer **Verringerung der Ergebnissicherheit** führen. Auf der anderen Seite können z. B. große Effekte eine Erhöhung der Sicherheit begründen.“



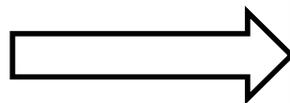
## (3 ↔ 4) Konsistenzbetrachtungen (WF)

Aussage		Anforderung		
		Anzahl der Studien	Ergebnissicherheit	Effekt
Beleg	↑↑	≥ 2	mehrheitlich <b>hoch</b>	„gleichgerichtet“ <sup>a</sup>
Hinweis	↑	≥ 2	mehrheitlich <b>mäßig</b>	„gleichgerichtet“ <sup>a</sup>
		<b>1</b>	<b>hoch</b>	statistisch <b>signifikant</b>
Anhaltspunkt ↗		≥ 2	mehrheitlich <b>gering</b>	„gleichgerichtet“ <sup>a</sup>
		<b>1</b>	<b>mäßig</b>	statistisch <b>signifikant</b>

a: **gleichgerichtet:**

- homogen: gepoolter Effekt statistisch **signifikant**
- heterogen: Effektrichtung zu > 80% dieselbe und Effekte zu > 50% statistisch **signifikant** (gemessen am Gewicht der Studien im REM)

## (4 ↔ 3) „Signifikanz“



**Ableitung von Belegen, Hinweisen, Anhaltspunkten für Nutzen/Schaden oder „nichts“.**

## Vergleiche

- Galantamin vs. Placebo (the
- Galantamin-Weiterbehandlung vs. Placebo
- Rivastigminpflaster 10cm vs. Placebo

Gemäß „Abschnitt 3.1.4  
Zusammenfassende Bewertung“:  
➤ **Welche Nutzensaussagen  
waren (nicht) möglich?**  
➤ **Warum?**

## Wesentliche Endpunkte

- **Kognitive Leistungsfähigkeit (ADAS-cog)**
- Aktivitäten des täglichen Lebens (*versch. Scores*)
- Unerwünschte Ereignisse
- (...)



\* Ergänzungsauftrag!

## Vergleiche

- Galantamin vs. Placebo (the  
(n = **8 RCT**)
- Galantamin-Weiterbehandlung  
vs. Placebo  
(n = **1 RCT**)
- Rivastigminpflaster 10cm v  
(n = **2 RCT**)

Gemäß „Abschnitt 3.1.4  
Zusammenfassende Bewertung“: ⇨

- 1. Welche Nutzensaussagen  
wären maximal möglich?**
- 2. An welchen Stellen führt die  
Methodik zur Auf- oder Abwertung  
der Ergebnissicherheit?**

## Wesentliche Endpunkte

- **Kognitive Leistungsfähigkeit (ADAS-cog)**
- **Aktivitäten des täglichen Lebens (versch. Scores)**
- **Unerwünschte Ereignisse**
- **(...)**

IQWiG-Bericht - Nr. 118

**Cholinesterasehemmer bei  
Alzheimer Demenz:  
Ergänzungsauftrag  
Rivastigmin-Pflaster und  
Galantamin**

**Abschlussbericht**

Auftrag: A29-07  
Version: 1.0  
Stand: 03.02.2012

\* Ergänzungsauftrag!

## Galantamin – Placebo

Zusammenfassung der Beleglage n = 8 RCT

Therapieziel	Galantamin <sup>a</sup> vs. Placebo
<b>Patientenrelevante Therapieziele</b>	
Aktivitäten des täglichen Lebens	⇔
Kognitive Leistungsfähigkeit	↑↑
Begleitende Psychopathologie	⇔
Gesundheitsbezogene Lebensqualität	keine Daten
Vollstationäre Pflege (Institutionalisierung)	keine Daten
Mortalität	(⇔)
<i>Unerwünschte Ereignisse</i>	
Studienabbrüche wegen UE	↓↓
Schwerwiegende UE	⇔

**B?**



„nichts“.

Beleg!

# AB 09-05 Cholinesterasehemmer / Alzheimer

## Riva-10cm – Placebo n = 2 RCT

Therapieziel	Riva (Pflaster, 10 cm <sup>2</sup> ) vs. Placebo
<b>Patientenrelevante Therapieziele</b>	
Aktivitäten des täglichen Lebens	⇔
Kognitive Leistungsfähigkeit	↑ / ↗ <sup>a</sup>
Begleitende Psychopathologie	⇔
Gesundheitsbezogene Lebensqualität	keine Daten
Vollstationäre Pflege (Institutionalisierung)	keine Daten
Mortalität	(⇔)
<i>Unerwünschte Ereignisse</i>	

**B?**

(1) VP (hoch)



(2) WF ?



Effekt war groß / signifikant  
→ Beurteilung des Einfluss' des  
Verzerrungspotenzials

# Beurteilung des Einflusses von Verzerrungspotenzial

Ziel: Problemorientierte und realitätsnahe Beurteilung des  
Einflusses von VP →  
**Aussagen maximaler Aussagesicherheit.**



- A09-05 (Rivastigmin bei Alzheimer): n = 2, VP hoch  
**Robustheitsprüfung**  
 VP: differenziell fehlende Werte in der Auswertung  
 Prüfung: Test nach konservativer Imputation (Sensitivitätsanalyse) ☹️
- D12-08 (Vemurafenib bei Malig. Melanom) n=1, VP hoch  
**Robustheitsabschätzung**  
 VP: Offene Studie; Patienten erhielten nach Therapieabbruch  
 andere Behandlungen  
 Prüfung: Sichtung der anderen Behandlungen, Überschätzung des  
 Effektes unwahrscheinlich 😊  
 → „Hinweis“ bei Gesamtüberleben

# AB 09-05 Cholinesterasehemmer / Alzheimer

## Riva-10cm – Placebo n = 2 RCT

Therapieziel	Riva (Pflaster, 10 cm <sup>2</sup> ) vs. Placebo
<b>Patientenrelevante Therapieziele</b>	
Aktivitäten des täglichen Lebens	⇔
Kognitive Leistungsfähigkeit	↑ / ↗ <sup>a</sup>
Begleitende Psychopathologie	⇔
Gesundheitsbezogene Lebensqualität	keine Daten
Vollstationäre Pflege (Institutionalisierung)	keine Daten
Mortalität	(⇔)
<i>Unerwünschte Ereignisse</i>	



(1) VP hoch

(2) WF ...

(3) Konsistenz ?

Unterschiedliche Effekte in Subgruppen (Alter)!

Ziel: **Aufdecken** von Effektmodifikation (EM)  
bzw. **unterschiedlichen Effekten in Teilpopulationen**  
gemäß IQWiG-Auftrag;  
Bei Hinweisen: Einfluss ggf. (nur) auf **Aussagesicherheit**

Methodik: **Regelmäßige** Durchführung a priori  
geplanter Subgruppenanalysen.



- A09-05 (Rivastigmin bei Alzheimer)

## Test auf Interaktion

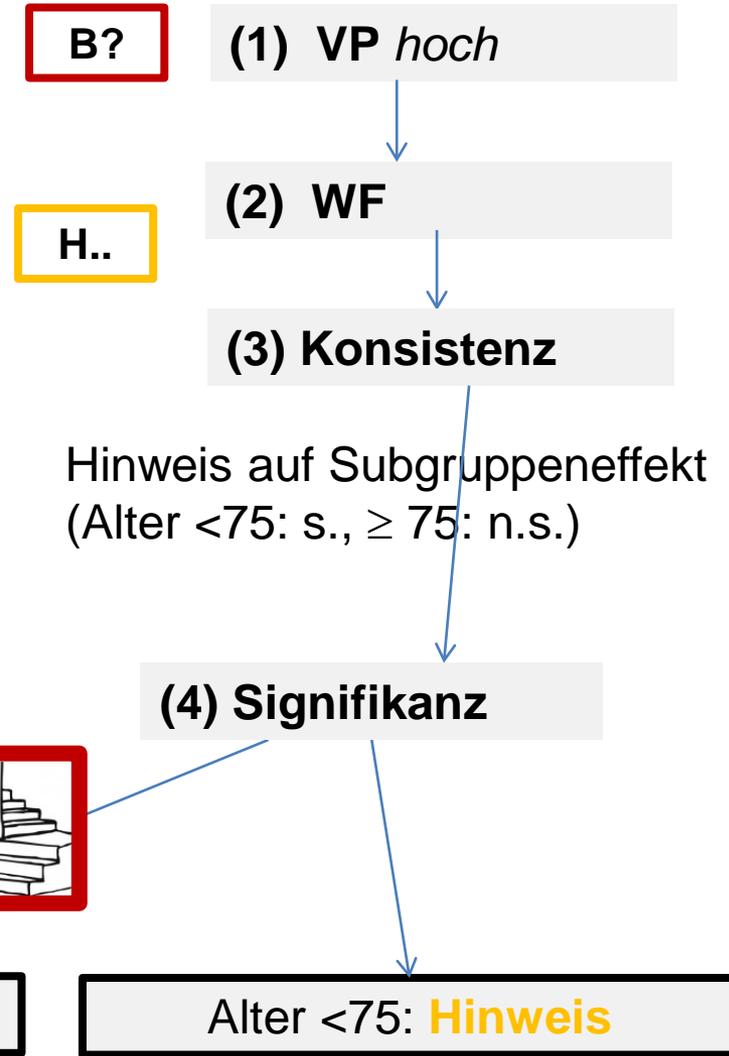
### **p<0,2: Hinweis auf EM**

- Gemeinsamen + Subgruppen-Schätzer darstellen
- Aussagen getrennt für Subgruppen
- bei Inkonsistenz (Gesamt- und Subgruppe)  
Aussagesicherheit für betroffene Subgruppe  
herabstufen (trifft hier zu)

# AB 09-05 Cholinesterasehemmer / Alzheimer

## Riva-10cm – Placebo n = 2 RCT

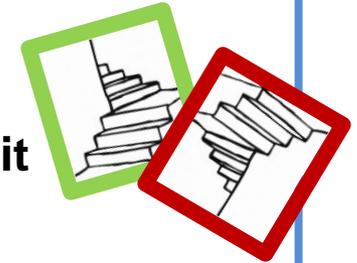
Therapieziel	Riva (Pflaster, 10 cm <sup>2</sup> ) vs. Placebo
<b>Patientenrelevante Therapieziele</b>	
Aktivitäten des täglichen Lebens	⇔
Kognitive Leistungsfähigkeit	
Begleitende Psychopathologie	⇔
Gesundheitsbezogene Lebensqualität	keine Daten
Vollstationäre Pflege (Institutionalisierung)	keine Daten
Mortalität	(⇔)
<i>Unerwünschte Ereignisse</i>	



# Subgruppenanalysen – Durchführung und Konsequenzen

Ziel: **Aufdecken** von Effektmodifikation (EM)  
bzw. **unterschiedlichen Effekten in Teilpopulationen**  
gemäß IQWiG-Auftrag;  
Bei Hinweisen: Einfluss ggf. (nur) auf **Aussagesicherheit**

Methodik: **Regelmäßige** Durchführung a priori geplanter  
Subgruppenanalysen.



## Test auf Interaktion: Positive Testausgänge

### **$p < 0,05$ : Beleg für EM**

#### **Darstellung und Aussagen:**

- Aussagen nur getrennt für Subgruppen
- nur Subgruppen-Schätzer darstellen

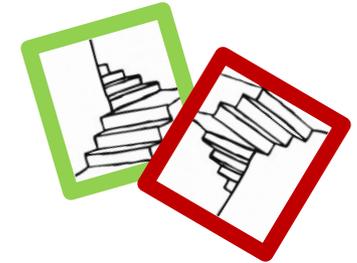
### **$0,05 \leq p < 0,2$ : Hinweis auf EM**

#### **Darstellung und Aussagen:**

- Aussagen getrennt für Subgruppen
- gemeinsamen + Subgruppen-Schätzer darstellen
- bei Inkonsistenz zwischen Gesamt- und Subgruppe ggf. Aussagesicherheit für Subgruppe herabstufen

**Nicht auffällige Tests** führen zu keiner spezifischen Darstellung (Darstellung des Gesamtschätzers).

# Subgruppenanalysen – Durchführung und Konsequenzen



- **Hinweis EM** (A09-05: Rivastigmin bei Alzheimer)  
→ Aussagesicherheit für ältere Patienten  
**herabgestuft** (Hinweis → Anhaltspunkt)
- **Beleg EM** (D11-25: Telaprevir bei Hepatitis C)  
für den Endpunkt „virologisches Ansprechen“  
→ **Nutzenbeleg ausschließlich für Subgruppe** (Ausgangsviruslast)
- **Beleg EM** (D12-05: Rilpivirin bei HIV-Infektion)  
für den Endpunkt „virologisches Ansprechen“  
→ **Nutzenbeleg ausschließlich für Subgruppe** (Männer)
- **kein (!) Beleg/ Hinweis auf EM** (D11-30: Apixaban nach Gelenk-OP)  
für den Endpunkt „Tiefe Beinvenenthrombose“ durch OP-Typ (Hüft-OP, Knie-OP),  
dem Vorgehen entsprechend → **Nutzenbeleg für die Gesamtpopulation**
- **„Ausnahme“: Erklären vorhandener Heterogenität** (D11-26: Eribulin bei  
Mamma-CA) durch **nicht präspezifizierte** Subgruppenanalyse  
(„Behandelbarkeit“ mit Alternativen)  
→ Ableiten von **Nutzen** (Überleben) **für eine Subgruppe möglich**

⇒ stringentes Verfahren

# AB 09-05 Cholinesterasehemmer / Alzheimer

## Galantamin-Weiterbehandlung - Placebo

n = 1 RCT

Therapieziel	Galantamin Weiterbehandlung nach Vorbehandlung
<b>Patientenrelevante Therapieziele</b>	
Aktivitäten des täglichen Lebens	(=)
Kognitive Leistungsfähigkeit	(/)
Begleitende Psychopathologie	keine Daten

**H?**

(1) VP (gering)



(2) WF (keine)



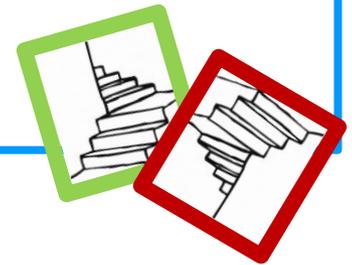
(3) Konsistenz ?

Mehrere Operationalisierungen  
des Endpunktes ADAS-Cog  
(einmal „üblich“, einmal ÜLZ-A)

# Konsistenz im Endpunkt („Weitere Faktoren“)

Ziel: Plausible endpunktbezogenen Interpretation der Ergebnisse  
(„Konsistenz“; Niveau-Überlegungen!)

Methodik: **Zusammenfassende Auswertung** aller erhobenen Daten  
zu einem Endpunkt, falls keine Priorisierung  
einzelner Operationalisierungen vorgesehen sind.



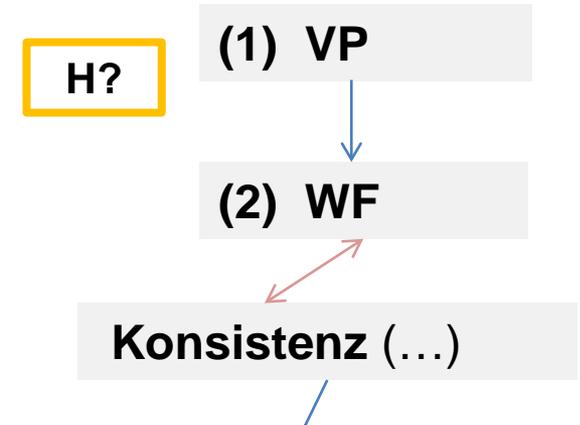
- A09-05 (Galantamin-Weiterbehandlung), ADAS-cog
  - 1. Operationalisierung: „üblich“; stetiger EP → n.s.
  - 2. Operationalisierung: Zeit bis Verschlechterung → signifikant
    - **Ergebnisse für alle Operationalisierungen darstellen**
    - **Gemeinsame Aussage für den Endpunkt ableiten**
    - **Bei Inkonsistenz ggf. Aussagesicherheit herabstufen**
- D11-26 (Eribulin Mamma-CA) n=1  
Endpunkt Überleben
  - 1. Auswertzeitpunkt für ÜLZ-A: → signifikant.
  - 2. Auswertzeitpunkt für ÜLZ-A → nicht signifikant  
(Ableitung eines Anhaltspunktes)

# AB 09-05 Cholinesterasehemmer / Alzheimer

## Galantamin-Weiterbehandlung - Placebo

n = 1 RCT

Therapieziel	Galantamin Weiterbehandlung nach Vorbehandlung
<b>Patientenrelevante Therapieziele</b>	
Aktivitäten des täglichen Lebens	(☹)
Kognitive Leistungsfähigkeit	(↗)
Begleitende Psychopathologie	keine Daten



Mehrere Operationalisierungen  
des Endpunktes ADAS-Cog

Veränderung stetig: nicht signifikant  
Zeit bis relevante Veränderung: signifikant.



Kognitive Leistungsfähigkeit:  
**Anhaltspunkt**

## ■ **Arzneimittel** (A09-05; Dossierbewertungen):

- Datenbasis: RCTs (teilweise n=1)

### 1. **Konsistenzbetrachtungen auf verschiedenen Ebenen**

- Homogenität der Effekte aus Studien
- Subgruppen
- Operationalisierung von Endpunkten

### 2. **Robustheitsprüfungen bzgl. des Verzerrungspotenzials**

- in ausgewählten Fällen

⇒ **Differenzierte Ableitung (Belegen, Hinweise, Anhaltspunkte, „nichts“...!)**

⇒ **Eindeutige Klassierung**

## ■ **Nichtmedikamentöse Interventionen** („Stammzellen“)

- Datenbasis: selten(er) RCTs; CCT; sogar non-CCTs
- Ergebnisse: primär Anhaltspunkte; Ausnahmen: Belege

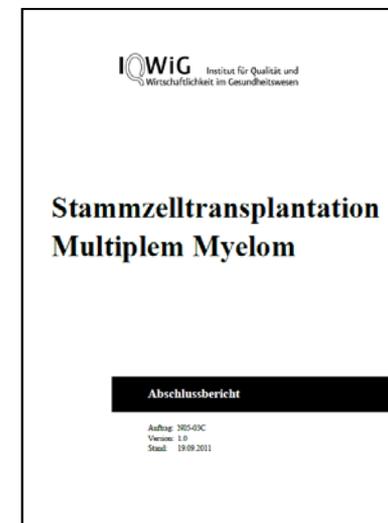
### 1. **Konsistenzbetrachtungen**

- Richtung der Effekte zwischen Studien

### 2. **Robustheitsprüfungen**

- Effektgrößen, Spezifität, Einfluss des Verzerrungspotenzials

⇒ **Methodik analog zur Nutzenbewertung von Arzneimitteln**



- **Regelhaftes Vorgehen maßgeblich für konsistente Aussagen**
- **„Handarbeit“ möglich**
- **neue Kategorie ermöglicht transparentes Vorgehen und differenzierte Abbildung der Aussagesicherheit**

## **Herausforderungen und geplante Weiterentwicklungen :**

- Zusammenfassung von Ergebnissen unterschiedlicher Ergebnissicherheit
- Beurteilung heterogener Ergebnisse in der Meta-Analyse („Gleichgerichtetheit“)
- „Indirektheit“

**“Science doesn’t know everything.”  
„Well, science *knows* that it doesn’t  
know everything. Otherwise it  
would stop.“ ...**

(Dara Ó Brian)

# Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

- Im Mediapark 8
- D-50670 Köln
  
- Telefon +49-221/3 56 85-0
- Telefax +49-221/3 56 85-1
  
- [info@iqwig.de](mailto:info@iqwig.de)
- [www.iqwig.de](http://www.iqwig.de)

