IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

The importance of non-inferiority testing in benefit assessments of medical interventions

Ralf Bender

Institute for Quality and Efficiency in Health Care
Dillenburger Str. 27 , 51105 Cologne

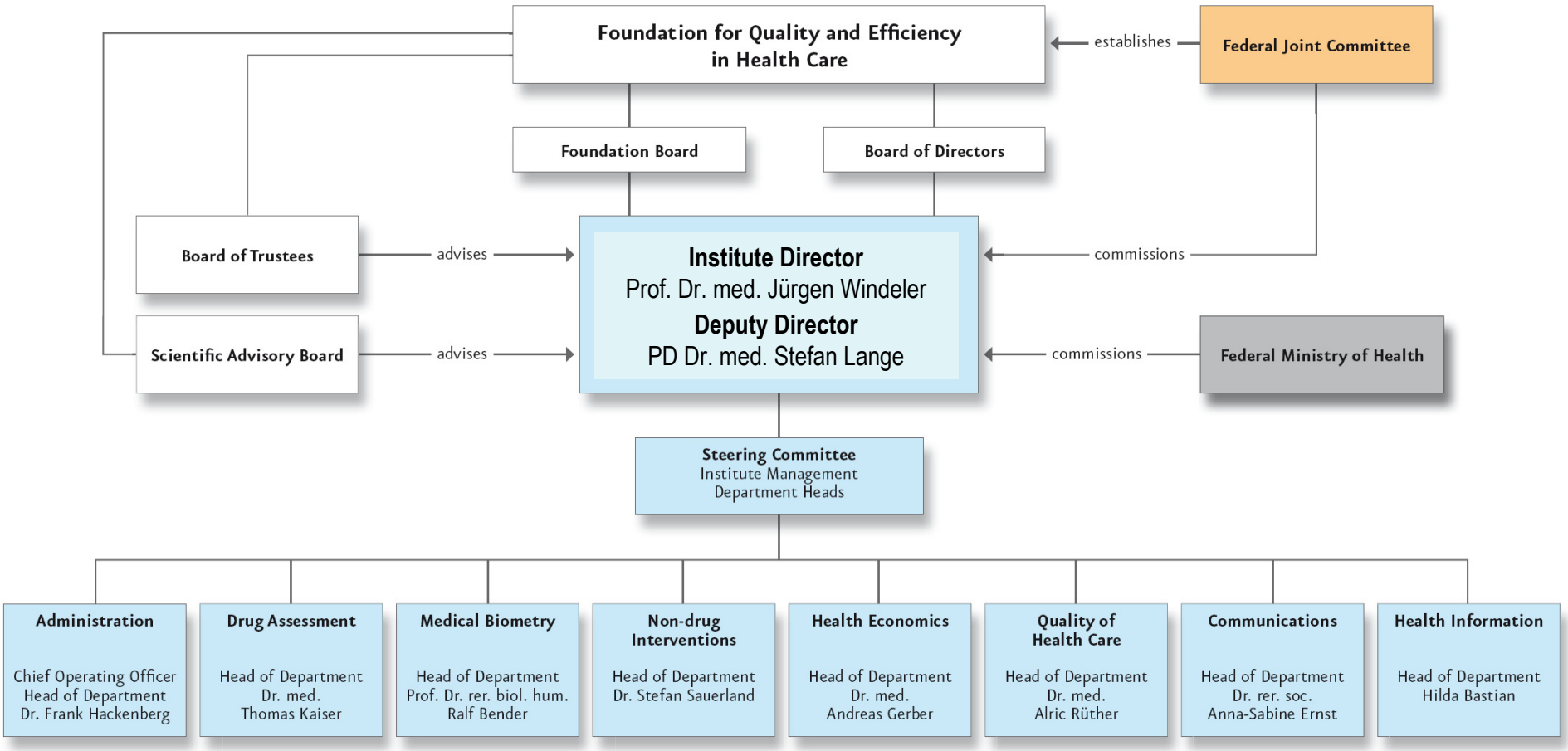
Agenda

- Institute for Quality and Efficiency in Health Care:
  - Legal responsibilities
  - Structure
  - Methods
- Non-inferiority and equivalence in IQWiG methods
- Examples:
  - Long acting insulin analogues in type 1 diabetes
  - Non-drug procedures for benign prostatic syndrome
- Conclusions
- References

The IQWiG

- § 139a Social Code Book (SGB V)
- Institute for Quality and Efficiency in Health Care
- (1) *„Der Gemeinsame Bundesausschuss nach § 91 gründet ein fachlich unabhängiges, rechtsfähiges, wissenschaftliches Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen und ist dessen Träger. Hierzu kann eine Stiftung des privaten Rechts errichtet werden.“*
- (2) […]

Structure


Foundation for Quality and Efficiency in Health Care
establishes
Federal Joint Committee
Foundation Board
Board of Directors
Board of Trustees
advises
Institute Director
Prof. Dr. med. Jürgen Windeler
Deputy Director
PD Dr. med. Stefan Lange
commissions
Scientific Advisory Board
advises
commissions
Federal Ministry of Health
Steering Committee
Institute Management
Department Heads
Administration
Chief Operating Officer
Head of Department
Dr. Frank Hackenberg
Drug Assessment
Head of Department
Dr. med.
Thomas Kaiser
Medical Biometry
Head of Department
Prof. Dr. rer. biol. hum.
Ralf Bender
Non-drug Interventions
Head of Department
Dr. Stefan Sauerland
Health Economics
Head of Department
Dr. med.
Andreas Gerber
Quality of Health Care
Head of Department
Dr. med.
Alric Rüther
Communications
Head of Department
Dr. rer. soc.
Anna-Sabine Ernst
Health Information
Head of Department
Hilda Bastian

Commission

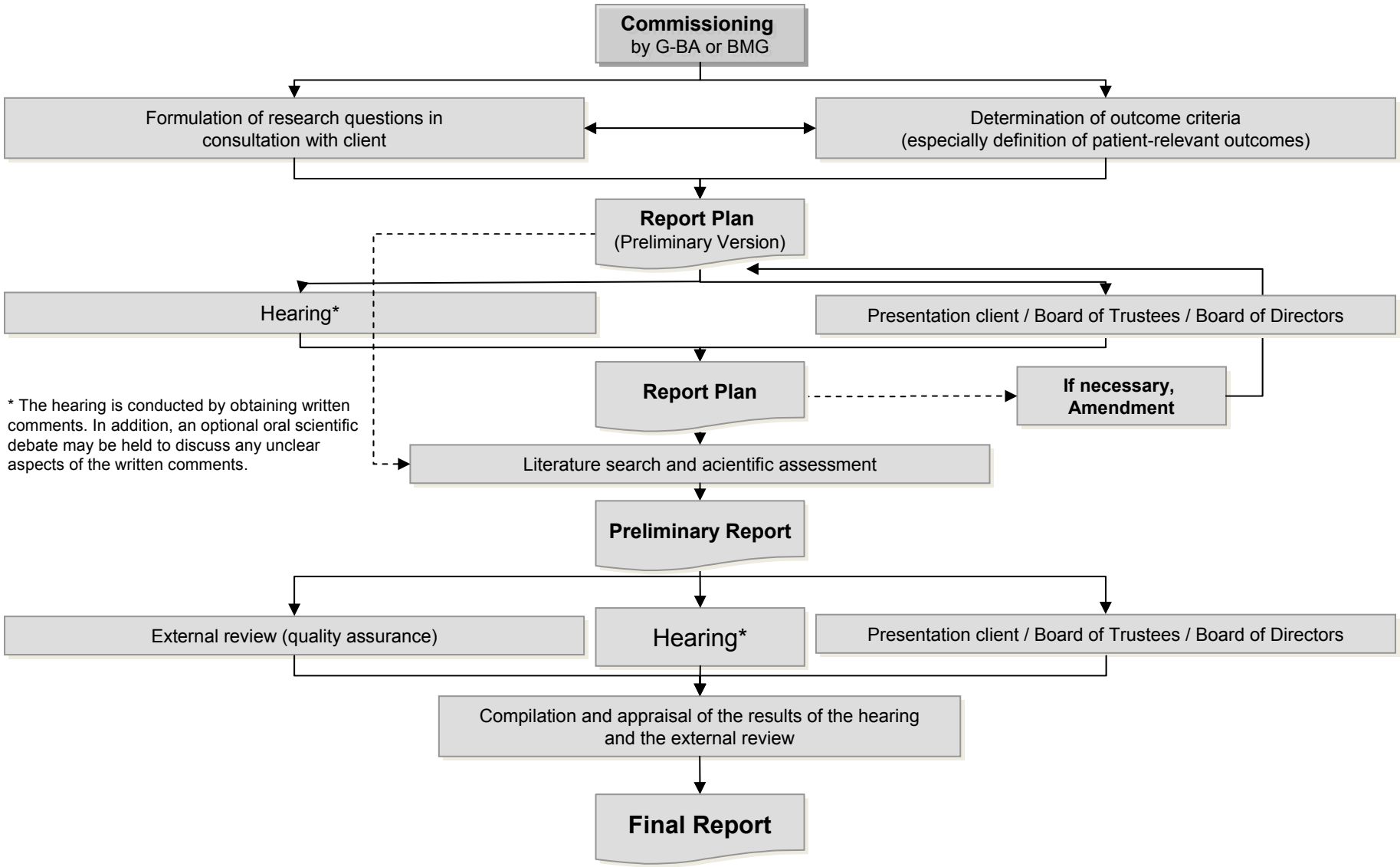
The Institute produces independent, evidence-based **reports**, e.g., on:

- Drugs
- Non-drug interventions
- Methods for diagnosing and screening
- Treatment guidelines
- Disease management programmes (DMPs)

In addition, IQWiG provides

- Health information for patients and the community

Production Procedure



Medical Biometry

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

Jobs:

- Biometrical advice and project support
- Scientific assessment of medical literature
- Statistical data analysis
- Development of statistical methods
- Publication of scientific knowledge
- Statistical education of IQWiG colleagues

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

General Methods

Version 3.0 of 27.05.2008

Non-inferiority

Chapter 3:
Assessment of benefit and harm of medical interventions

Section 3.1.4 *Summarizing assessment*

… exactly one of the five following evaluating conclusions is first drawn for each predefined patient-relevant outcome on the basis of the analysed scientific data available:

1. Proof of a(n) (additional) benefit or harm exists.
2. Indications of a(n) (additional) benefit or harm exist.
3. Proof of the lack of a(n) (additional) benefit or harm exists.
4. Indications of the lack of a(n) (additional) benefit or harm exist.
5. No proof and no indication of a(n) (additional) benefit or harm exist.

…

Well-founded definitions of irrelevance ranges are the precondition for indications or proof of the lack of a(n) (additional) benefit or harm (see Section 6.4.5).

…

Non-inferiority

Chapter 6: **General methodological aspects**

Section 6.4.5 *Demonstration of equivalence*

One of the most common serious errors in the interpretation of medical data is to rate the non-significant result of a traditional significance test as evidence that the null hypothesis is true [9]. To demonstrate "equivalence", methods to test equivalence hypotheses need to be applied [216].

…

In addition, in equivalence studies the equivalence range must be clearly defined. This range can be two-sided, resulting in an equivalence interval, or one-sided in terms of an "at most irrelevant difference" or "at most irrelevant inferiority". The latter is referred to as a "non-inferiority hypothesis" [82,200,303].

…

Specifically developed methods should be applied to analyse data from equivalence studies. The "confidence interval inclusion method" is a frequently used technique. …

Chapter 6: **General methodological aspects**

Section 6.4.5 *Demonstration of equivalence*

…

Compared with superiority studies, equivalence studies show specific methodological problems. On the one hand, it is often difficult to provide meaningful definitions of equivalence ranges [243]; on the other hand, the usual study design criteria, such as randomisation and blinding, no longer sufficiently protect from bias [332]. Even without knowledge of the treatment group, it is possible, for example, to shift the treatment differences to zero and hence in the direction of the desired alternative hypothesis. Moreover, the ITT principle should be applied carefully, as its inappropriate use may falsely indicate equivalence [216]. For this reason, particular caution is necessary in the evaluation of equivalence studies.

Example A05-01

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

Long-acting insulin analogues (LAIAs) in the treatment of diabetes mellitus type 1 (only adults considered here)

- Commission by G-BA
- Aim: Benefit assessment of LAIAs
- 14 endpoints (mortality, morbidity, quality of life)
- 12 relevant studies (some are non-inferiority trials):
  - 6 studies glargine vs. human insulin (1 non-inferiority trial)
  - 4 studies detemir vs. human insulin (all non-inferiority trials)
  - 2 studies glargine vs. detemir (all non-inferiority trials)
- Conclusions:
  - No proof of additional benefit for glargine vs. human insulin
  - No proof of additional benefit for detemir vs. human insulin
  - No proof of additional benefit for glargine vs. detemir

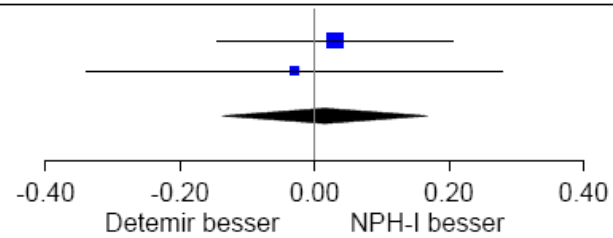
Example A05-01

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

Comparison detemir vs. human insulin with once- to twice-daily administration (endpoint HbA1c change):

Detemir vs. NPH-Insulin
HbA1c-Änderung
Modell mit zufälligen Effekten - DerSimonian und Laird

Studie	Effekt	SE	Effekt (95%-KI)	Gewichtung	Effekt	95%-KI
1476	0.03	0.09		75.8	0.03	[-0.15, 0.21]
1582	-0.03	0.16		24.2	-0.03	[-0.34, 0.28]
Gesamt				100.0	0.02	[-0.14, 0.17]



Heterogenität: Q=0.11, df=1, p=0.741, I²=0%
Gesamteffekt: Z Score=0.20, p=0.842, Tau=0

- Both studies showed non-inferiority (margin 0.4)
- CI of pooled effect measure also would show non-inferiority
- However, this plays nor role for IQWiG conclusion
- Aim was to show ***additional benefit*** regarding patient relevant endpoints (IQWiG, 2010)

Example N04-01

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

Non-drug local procedures in the treatment of benign prostatic hyperplasia (BPH)

- Commission by G-BA
- Aim: Comparison of non-drug local procedures for BPH
- 21 main procedures (6 “standard”, 15 “newer”)
- 6 endpoints (symptoms, hospitalisation, QoL, and others)
- 56 relevant studies (55 RCTs, 1 CCT)
- Conclusions:
  - None of the 15 “newer” procedures showed improved or **equivalent** symptom relief compared to standard
  - Some procedures have the advantage of shorter hospital stays
  - One procedure can be performed on an outpatient basis

Example N04-01

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

Conclusions of preliminary report (version 1.0):

- Indications of benefit for standard compared to some of the newer procedures regarding symptoms
- Longer hospital stay for standard procedures
- (Further conclusions regarding adverse events.)

Valid argument of hearing:

- Usual definition of additional benefit regarding symptoms is critical
- Less invasive procedures have a clear additional benefit if they are non-inferior to standard regarding symptoms
- Consequence: Establishment of irrelevance range

Example N04-01

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

Establishment of irrelevance range:

- None of the 56 studies was planned as equivalence or non-inferiority trial
- Based on Lange & Freitag (2005) an irrelevance margin of 0.25 SD regarding mean difference was chosen
- Sensitivity analyses with margin of 0.5 SD
- Range of SDs of symptom scores: 7 to 10
- Median SD 8.5 $\Rightarrow$
  - 0.25 SD $\approx$ 2.1 score points
  - 0.5 SD $\approx$ 4.3 score points
- For sample size estimation 0.4 to 0.8 SD was used
- Non-inferiority if UL of CI for Hedges g below 0.25
- Superiority if UL of CI for Hedges g below 0

Example N04-01

Results regarding symptoms:

- None of the 15 “newer” procedures showed non-inferior symptom relief compared to standard
- Even if the “large” margin 0.5 SD was used 14 of the “newer” procedures are not non-inferior
- One procedure showed non-inferiority if the “large” margin 0.5 SD was used; this should be investigated in further trials, see IQWiG (2008)

Conclusions

- Non-inferiority and equivalence are important in benefit assessments of medical interventions
- However, non-inferiority and equivalence tests are rarely performed in benefit assessments
- It is impossible to define irrelevance ranges in systematic reviews *before* data realisation
- It is difficult to agree upon irrelevance margins if there are conflicts of interest
- Best approach is to use widely accepted irrelevance margins for a specific endpoints
- If no accepted irrelevance margin is available a fixed margin regarding SMD should be used complemented by sensitivity analyses

References

- IQWiG (2008): *General Methods*. Version 3.0 of 27.05.2008. Institute for Quality and Efficiency in Health Care, Cologne.
- IQWiG (2008): *Non-drug local procedures in the treatment of benign prostatic hyperplasia*. Final Report N04-01, Version 1.0 [in German]. Institute for Quality and Efficiency in Health Care, Cologne.
- IQWiG (2010): *Long-acting insulin analogues in the treatment of diabetes mellitus type 1*. Final Report A05-01, Version 1.0 [in German]. Institute for Quality and Efficiency in Health Care, Cologne.
- Lange, S. & Freitag, G. (2005): Choice of delta: Requirements and reality – results of a systematic review. *Biometrical Journal* **47**, 12-27.

Institute for Quality and Efficiency in Health Care

- Dillenburger Straße 27
- D-51105 Köln

- Phone +49 221 / 35685-0
- Fax +49 221 / 35685-1

- info@iqwig.de
- www.iqwig.de