Is “15% of the scale range” Universally Applicable to Define “MID” and Clinical Relevance of Patient-Reported Treatment Benefits?

Group Discussion with the ISPOR Clinical Outcome Assessment Special Interest Group (COA SIG)

SPEAKERS

Olivier Chassany, MD, PhD

Greater Paris University Hospitals (AP-HP), Paris, France

Katja Rudell, PhD, MSc

Parexel International, Uxbridge, London UK

James Shaw, PhD, PharmD, MPH

Bristol-Myers Squibb, Lawrenceville, NJ, USA

Academic perspective

Prof. Olivier Chassany
Patient-Report Outcomes Research (UMR 1123), Université de Paris, Inserm
Health Economics Clinical Trial Unit (URC-ECO), Hopital Hotel-Dieu, AP-HP, Paris

Declaration of recent interests

- Sanofi, Pfizer, Gilead, BMS, IQVIA, Boehringer

Example of a dossier submitted to the Federal Joint Committee (G-BA), using the usual MID (i.e. 5 points for PCS and MCS of the SF-36)

Health-related quality of life

Endpoint	Secukinumab		Adalimumab		Secukinumab vs Adalimumab
	N	Patients with event n (%)	N	Patients with event n (%)	RR [95 % CI]; p value[b]
SF-36					
Improvement of ≥ 5 points[2]					
Mental Component Score (MCS, improvement of 5 points)	110	68.5 (62.3)	101	45.3 (44.9)	1.39 [1.06; 1.83]; 0.018
PhysicalComponent Score (PCS, improvement of ≥ 5 points)	110	66.9 (60.8)	101	62.1 (61.5)	0.99 [0.79; 1.24]; 0.929

Diff: 23,2%

Resolution of the Federal Joint Committee (G-BA) on an Amendment of the Pharmaceuticals Directive (AM-RL): Annex XII – Benefit Assessment of Medicinal Products with New Active Ingredients According to Section 35a SGB V Secukinumab (Reassessment on the Basis of New Scientific Findings (Psoriatic Arthritis)) - 18 February 2021

German HTA agency IQWiG (Institute for Quality and Efficiency in Health Care) updated their general methods paper on health benefit assessment (05 Nov 20)

- Pharmaceutical manufacturers must submit benefit dossiers to facilitate HTA appraisals to G-BA (Federal Joint Committee) in Germany for new pharmaceutical, new combination or new label.
- For all non-orphan drugs the G-BA delegates the benefit assessment to the IQWiG. The IQWiG conducts the dossier evaluation using the criteria set out in the IQWiG general methods paper.
- The entire HTA is based on patient relevant endpoints, which include Clinical Outcomes Assessments and Patient Reported Outcomes (PRO) . The final decision on the added benefit is made by the G-BA.

https://www.iqwig.de/methoden/general-methods_version-6-0.pdf?rev=194070
https://www.g-ba.de/english/benefitassessment/

German HTA agency IQWiG (Institute for Quality and Efficiency in Health Care) updated their general methods paper on health benefit assessment (05 Nov 20)

- A within-patient relevance threshold was established to allow IQWiG make quick decisions about benefits of therapies that were using PRO rating scales.
- Based on a systematic literature review that covered 8 therapy areas*, the overall threshold of treatment benefits ranged between +1-38% change
- The new rule is to present results : % of patients who improve their scores by at least 15% of the width of an overall scale (e.g. 15 points on a 0-100 scale).
  - Patients exceeding that threshold are considered to perceive a meaningful change (responder)

** rheumatism, orthopaedics, paediatrics, fatigue, oncology, cardiovascular, COPD, urogenital, musculoskeletal diseases and miscellaneous diseases*

https://www.iqwig.de/methoden/general-methods_version-6-0.pdf?rev=194070
https://www.g-ba.de/english/benefitassessment/

IQWiG skepticism regarding minimal important difference (MID) one of the standard threshold for assessing benefit

IQWiG criticizes MIDs:

- validation not scientific sound
- high variability
- Lack of standard to assess quality of MID
- Lack of quality in MID reporting
- MIDs were not accepted anymore by IQWiG, but often by the German Regulatory Agency that assessed the efficacy and safety of the product (G-BA) due to consistency reasons.

https://www.iqwig.de/en/presse/press-releases/press-releases-detailpage_27520.html

IQWiG ABOUT US PROJECTS PARTICIPATION EVEN

PRESS > PRESS RELEASES > CLINICAL RELEVANCE OF PATIENT-REPORTED OUTCOMES: NEW THRESHOLD PROVEN FEASIBLE IN PRACTICE

Feb 18, 2021

Clinical relevance of patient-reported outcomes: new threshold proven feasible in practice

Current examples prove the feasibility of the 15 percent threshold newly defined by IQWiG for the acceptance of responder analyses in early benefit assessments of drugs.

There were concerns that a strict 15-percent threshold for **responder analyses** would not be applicable to all scales, making it more difficult to prove added benefit. These concerns were unfounded.

The addenda published today show that the **new threshold is feasible in practice.**

Katrin Nink, IQWiG-Ressort Arzneimittelbewertung 02'2021

In order to show the clinical relevance of a difference between two treatment alternatives, in recent years, the manufacturer dossiers submitted in early benefit assessments of new drugs have increasingly contained responder analyses for patient-relevant outcomes. In such analyses, it is investigated whether the proportion of patients experiencing a noticeable change in the respective outcome differs between the two treatment groups in a study. This involves information on health-related quality of life or on individual symptoms such as pain or itching, which patients recorded with the help of scales in questionnaires. But what difference makes a change relevant for the individual? That is, at what threshold can a response to an intervention be derived for the patient, so that, for example, the difference in the response rates of two groups can be used as an effect measure for early benefit assessments?

IQWiG skepticism regarding minimal important difference (MID) one of the standard threshold for assessing benefit

Discussion (June 2020) with stakeholders:

- IQWiG explained their 15% approach, which is not meant to be a MID but a response threshold, “where we are sufficiently sure that we represent a change that is noticeable for patients”.

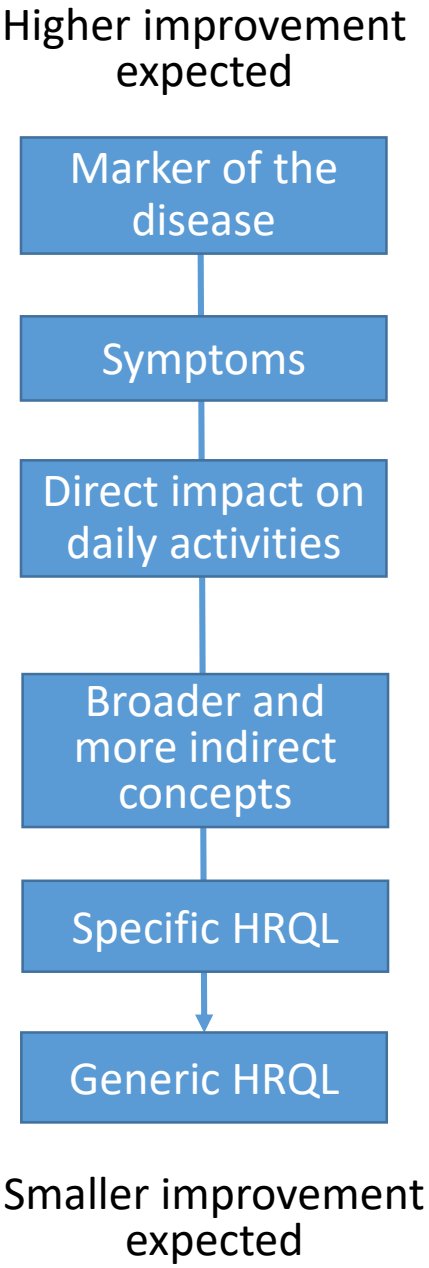
This is the nub of matter :

- “noticeable” is exactly what means the MID, a minimal change that patients are able to notice anchored to a patient global impression of change
- The challenge is to decide what level of change is large enough to define a patient as responder

A universally response threshold of 15% of the scale range is appealing but too simple

A unique 15% of change cannot fit all the responder definitions depending on the PRO endpoint used

- Responder definition (i.e. relevant change) depends of the concept captured by the tool
- 15% of scale range rule:
  - May not be large enough for primary symptoms and for endpoints which measure direct impact on daily life activities (e.g., impact of pyrosis improvement on diet)
  - Is likely to be not achievable for more indirect PRO such as Health-Related Quality of life, for which the improvement is smaller, especially, when the improvement on the clinical primary endpoint is already small.
- **We cannot expect an improvement larger on HRQL scores than on the direct marker of the evolution of the disease.**



Applying the 15% threshold of the scale range

Health-related quality of life

Endpoint	Secukinumab		Adalimumab		Secukinumab vs Adalimumab
	N	Patients with event n (%)	N	Patients with event n (%)	RR [95 % CI]; p value[b]
SF-36					
Improvement by 15 % of the range of the scale[2]					
Mental Component Score (MCS, improvement of 9.6 points)	110	46.1 (41.9)	101	28.8 (28.5)	1.47 [0.99; 2.19]; 0.055
PhysicalComponent Score (PCS, improvement of ≥ 9.4 points)	110	42.8 (39.0)	101	37.9 (37.5)	1.04 [0.73; 1.49]; 0.834
Improvement of ≥ 5 points[2]					
Mental Component Score (MCS, improvement of 5 points)	110	68.5 (62.3)	101	45.3 (44.9)	1.39 [1.06; 1.83]; 0.018
PhysicalComponent Score (PCS, improvement of ≥ 5 points)	110	66.9 (60.8)	101	62.1 (61.5)	0.99 [0.79; 1.24]; 0.929

Diff: 17,3%

Diff: 23,2%

Resolution of the Federal Joint Committee (G-BA) on an Amendment of the Pharmaceuticals Directive (AM-RL): Annex XII – Benefit Assessment of Medicinal Products with New Active Ingredients According to Section 35a SGB V Secukinumab (Reassessment on the Basis of New Scientific Findings (Psoriatic Arthritis)) - 18 February 2021

Applying the 15% of the scale range threshold on a generic heath status scale is counter productive

Very unlikely to show a difference between groups based on a such high responder definition (i.e. 15 mm of EQ-5D 0-100 mm VAS scale) unless the treatment has a dramatic benefit

Morbidity

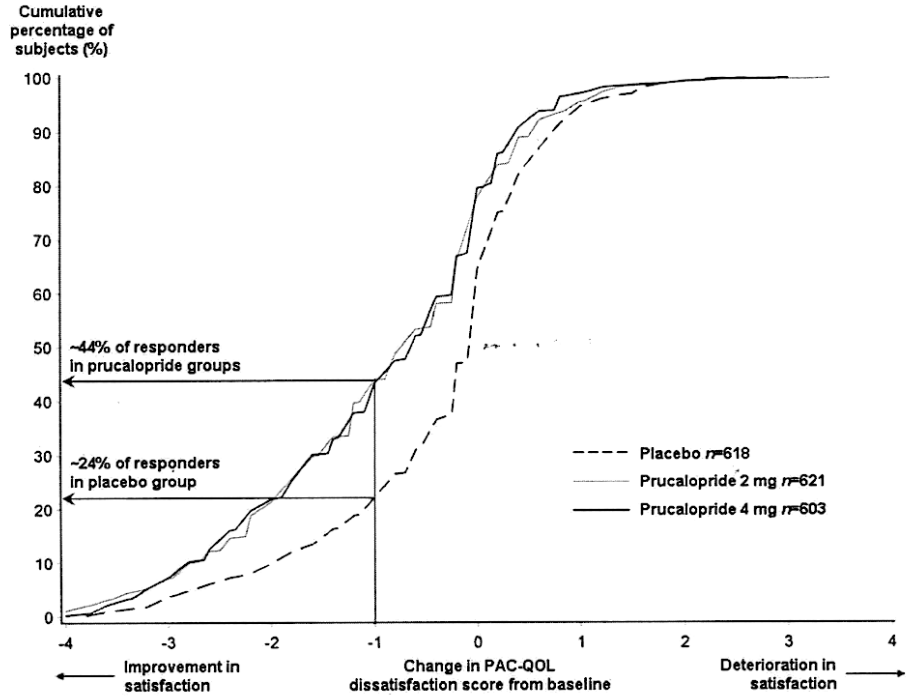
Endpoint	Secukinumab		Adalimumab		Secukinumab vs adalimumab
	N	Patients with event n (%)[a]	N	Patients with event n (%)[a]	RR [95 % CI]; p value[b]
Health status (EQ-5D VAS, improvement of ≥ 15 mm, ≙ 15 % of the range of the scale)	110	58.8 (53.5)	101	60.2 (59.6)	0.90 [0.70; 1.15]; 0.388

Resolution of the Federal Joint Committee (G-BA) on an Amendment of the Pharmaceuticals Directive (AM-RL): Annex XII – Benefit Assessment of Medicinal Products with New Active Ingredients According to Section 35a SGB V Secukinumab (Reassessment on the Basis of New Scientific Findings (Psoriatic Arthritis)) - 18 February 2021

No scientific reason why the responder definition should be different for drug approval regulators and for HTA assessors

This “universal” 15% of the scale range rule applied for HTA assessment, contradicts the EMA and FDA guidelines which recommend sensitivity analyses using several definitions of responders, and/or e.g. to present results of change as a cumulative distribution function (CDF) across the range scale, which is more informative than a single responder definition.

Linaclotide vs placebo	12-wk A	12-wk B	26-wk B
Abdominal pain responder (30%)	55 vs 42% **Δ = 13%**	54 vs 39% **Δ = 15%**	54 vs 36% **Δ = 18%**
IBS relief responder	37 vs 19% **Δ = 16%**	39 vs 17% **Δ = 16%**	37 vs 17% **Δ = 20%**
Abd pain responder (40%)	48 vs 34%	**← Sensitivity analyses**	
Abd pain responder (50%)	39 vs 26%		
Sustained • Pain responder • IBS relief responder	53 vs 41% 34 vs 18%		



European public assessment report (EPAR) - http://www.ema.europa.eu

Conclusion - The 15% of the range scale rule cannot be the universal solution for interpreting PRO results

It may be easier to base a decision on a single binary (> 15%) presentation of results, but it reduces considerably the information usually needed to review a dossier with PRO/COA data, being for MA or for HTA

The added value of PROs is also based on the review:

- Evolution of scores over time (if repeated)
- Consistency :
  - Across different PROs (symptoms, HRQL, work productivity, satisfaction…)
  - With the other endpoints of study
  - Across studies

The IQWIG rule:

- is not tailored to each of the concepts (direct vs indirect)
- is just one responder definition among several possible
- does not resolve the issue of whether the difference in responders between groups (comparative trial or indirect comparison) is relevant or not