

Stellungnahme der Ecker + Ecker GmbH zur Änderungen der Modulvorlage

Autoren:

Dr. Christof Ecker

Dr. Thomas Ecker

Dr. Claudia Pütz

Dr. Johannes Klaus

Dr. Stephanie Stengel

Dr. Meike von Wulffen

Einleitung

Die Ecker + Ecker GmbH unterstützt pharmazeutische Unternehmen regelmäßig im Verfahren der frühen Nutzenbewertung. In dieser Kompetenz nimmt die Ecker + Ecker GmbH Stellung zur geplanten Änderung der Modulvorlage, zu der vom G-BA mit Beschluss vom 17. Juni 2021 ein Stellungnahmeverfahren eröffnet wurde [1].

Diese geplante Änderung soll die vom IQWiG in seinem Methodenpapier vorgeschlagene Vorgehensweise zur Interpretation von komplexen Skalen in die Modulvorlage übernehmen. Das Vorgehen sieht vor, nur noch *minimal important clinical differences* (MCID) als Response-Kriterien anzuwenden, wenn diese oberhalb von 15 % der Skalenspannweite liegen (im Folgenden „15 %-Kriterium“). Liegt die für das Instrument validierte MCID unterhalb dieser Grenze, kann entweder auf standardisierte Mittelwertdifferenzen zurückgegriffen werden oder eine Analyse mit einem Response-Kriterium von genau 15 % der Skalenspannweite durchgeführt werden [2].

Die Ecker + Ecker GmbH sieht dieses Vorgehen kritisch. Schon der Grundgedanke, ein klinisches Relevanzkriterium aus der Skalenspannweite eines Instrumentes abzuleiten, ist nicht sachgerecht. Insbesondere ist jedoch die Herleitung des 15 %-Kriteriums nach wie vor fragwürdig. Darüber hinaus stellt die mit der allgemeinen Implementierung des 15 %-Kriteriums verbundene Ablehnung etablierter und für die Messung spürbarer Änderungen entwickelter MCID eine Abwertung der Patientenperspektive in der Interpretation patientenberichteter Endpunkte dar. Schließlich bleibt die Praktikabilität eines pauschalen, auf alle Skalen und Instrumente angewandten Response-Kriteriums, fraglich.

1 Die Herleitung des 15 %-Kriteriums bleibt fragwürdig

Auch nach der erfolgten Würdigung der Stellungnahmen zur Änderung des Methodenpapiers des IQWiG bleibt festzustellen, dass die Herleitung des 15 %-Kriteriums nicht evidenzbasiert erfolgte und auf einer zweifelhaften Grundlage beruht. Nach der initialen Kritik durch viele Stellungnehmer hat das IQWiG zwar die erfolgte Literatursuche zu bislang publizierten MCID detaillierter dargestellt. Jedoch erscheint das Vorgehen immer noch willkürlich und ist nicht nachvollziehbar.

So werden sehr niedrige MCID mit der Begründung, sie seien „unrealistisch klein“ nicht mit einbezogen, große MCID jedoch schon. Dabei stellt das IQWiG auch für große MCID fest, dass es „die Frage, ob die Werte am oberen Rand des Spektrums verlässlicher sind, [...] aufgrund fehlender akzeptierter Standards zur Qualitätsbewertung, wie auch fehlender Berichtsqualität von MID-Studien, nicht beantworten“ kann [3]. Warum daher große MCID nicht ebenfalls aus der Betrachtung ausgeschlossen werden, bleibt offen. Darüber hinaus hat das IQWiG für den Anspruch, dass 15 %-Kriterium würde hinreichend sicher unabhängig von Skala und Kontext für den Patienten relevante Veränderung abbilden, keinerlei Evidenz vorgelegt.

Ohnehin beruht die Herleitung des 15%-Kriteriums auf einer widersprüchlichen Argumentation. Denn die Begründung dafür, ein neues Vorgehen zur Bestimmung einer klinischen Relevanzschwelle zu entwickeln, ist die von IQWiG und G-BA konstatierte unklare Qualität bisheriger MCID. So schreibt der G-BA in den Tragenden Gründen zum Beschluss für dieses Stellungnahmeverfahren:

„Demnach zeigen systematische Zusammenstellungen empirisch ermittelter MIDs, dass zu einzelnen Instrumenten häufig eine Vielzahl von MIDs publiziert werden, die innerhalb eines Erhebungsinstruments große Spannweiten haben können. Ursächlich hierfür können unter anderem die in den Studien eingesetzten unterschiedliche Anker, Beobachtungsperioden oder analytische Methoden sein. Gleichzeitig ist eine anhand methodischer Qualitätskriterien begründete Auswahl empirisch ermittelter MIDs für die Nutzenbewertung derzeit nicht zu treffen“ [4]

Wenn die bisherigen MCID in ihrer Qualität jedoch nicht beurteilbar sind, können sie schlechterdings nicht für eine Herleitung eines übergeordneten Kriteriums herangezogen werden. Daher ist die Herleitung des 15%-Kriteriums durch eine mehr oder minder systematische Zusammenschau bisheriger MCID nicht nachvollziehbar. Genauso gut könnte man im Umkehrschluss folgern, dass die zur Ableitung dieses Kriteriums herangezogenen MCID valide sein müssen. Dann aber könnten sie auch direkt als klinisches Relevanzkriterium herangezogen werden.

Die Inkonsistenz dieses Vorgehens spiegelt sich auch in der Tatsache wider, dass MCID, die größer als 15% der Skalenspannweite sind, in der frühen Nutzenbewertung herangezogen werden sollen. Die Anerkennung einer MCID hängt dadurch von ihrem Ausmaß ab und nicht von der zu ihrer Herleitung verwendeten Methodik. Es ist jedoch die Qualität der Methodik, die über die Validität einer MCID entscheidet. Dabei unterscheidet sich die Qualität der zur Ermittlung von MCID angewandten Methodik nicht grundsätzlich zwischen MCID, die unterhalb von 15% der Skalenspannweite liegen und solchen, die oberhalb dieser Grenze liegen. Wenn MCID über dieser Grenze als valide betrachtet werden können, gilt dies für MCID unterhalb dieser Schwelle natürlich nicht minder.

2 Die Bedeutung von patientenberichteten Endpunkten (PRO) wird durch dieses Vorgehen in Frage gestellt

Für die Interpretation von patientenberichteten Endpunkten stellen MCID einen zentralen Bestandteil dar. Sie dienen dazu, eine für den Patienten spürbare und wichtige Veränderung sensitiv von trivialen Änderungen zu trennen und sollen die Patientenperspektive in die Interpretation des Instrumentes einbinden. Aus diesem Grund werden sie nicht willkürlich gewählt, sondern idealerweise anhand von klinischen Kriterien abgeleitet, welche die Patientenperspektive berücksichtigen sollen und durchaus zu je nach Indikation unterschiedlichen MCID führen können. Beispiele für solche klinischen Kriterien sind patientenberichtete Anker. Das skalen- und indikationsübergreifende 15 %-Kriterium beruht auf keiner derartigen klinischen Validierung und negiert daher die Patientenperspektive für die Interpretation von patientenberichteten Endpunkten.

So ignoriert das 15%-Kriterium Veränderungen der Lebensqualität oder anderer patientenrelevanter Parameter, die unterhalb der 15%-Schwelle liegen, aber für die Patienten spürbar und wichtig sind. Umgekehrt kann das Kriterium zur Anwendung einer zu niedrigen Relevanzschwelle führen, die fälschlicherweise einen Patientennutzen suggeriert. So würde die Anwendung auf die *Symptom Severity Scale* des *Uterine Fibroid Symptom and Quality of Life Questionnaire* (UFS-QoL) zu einer Relevanzschwelle von 15 Punkten führen, obwohl das IQWiG davon ausgeht, dass selbst 17 Punkte keine für die Patienten spürbare Änderung bedeuten [5]. Zudem werden Analysen, welche die Stabilität eines erreichten Gesundheitszustandes nachweisen wollen (beispielsweise die Zeit bis zur Verschlechterung der Lebensqualität) durch die Anwendung des 15%-Kriteriums erschwert, da kleinere Veränderung, die für den Patienten eine relevante Verschlechterung darstellen könnten, ignoriert werden.

Aufgrund der hohen Relevanz, welche der Lebensqualität in der frühen Nutzenbewertung zukommen soll, ist es nicht nachvollziehbar, warum für die Patienten spürbare Effekte ignoriert werden, nur weil sie nicht die Grenze von 15 % der Skalenspannweite erreichen, und Änderungen oberhalb von 15 % der Skalenspannweite pauschal als relevant eingestuft werden sollen. Durch die Abkopplung des Relevanzkriteriums von für die Patienten spürbaren Kriterien besteht die Gefahr, einen patientenrelevanten Nutzen nicht zu erkennen oder einen für die Patienten wichtigen Schaden nicht als solchen zu detektieren. Der Wert patientenberichteter Endpunkte wird durch ein solches Vorgehen in Frage gestellt.

3 Die Praktikabilität des 15%-Kriteriums bleibt fraglich

Das 15%-Kriterium soll als universell anwendbares Kriterium in die Dossievorlage aufgenommen werden. Doch schon jetzt zeichnet sich ab, dass das Kriterium nicht auf alle Skalen anwendbar ist. So lässt sich das 15 %-Kriterium nicht sinnvoll auf die theoretisch mögliche Skalenspannweite der Summenskalen des SF-36 anwenden, da in der Praxis die möglichen Maximalwerte nicht erreicht werden. Auch das IQWiG schlägt hier die Anwendung des Kriteriums auf die in einer Normstichprobe ermittelte Skalenspannweite vor, also 10 Punkte [6]. Das Problem kann aber auch für andere Instrumente bestehen und es bleibt offen, anhand welcher Kriterien entschieden werden soll, auf welche Skalenspannweite das 15 %-Kriterium angewendet werden soll.

4 Schlussfolgerung

Zusammenfassend erscheint uns der Nutzen des 15 %-Kriteriums fragwürdig, da den Vorteilen (Einfachheit) erhebliche Nachteile gegenüberstehen (u.a. methodische Fragwürdigkeit, Verlust an Sensitivität, unklare Anwendbarkeit auf alle Skalen). Zudem ist das Kriterium unnötig, da es dem G-BA schon jetzt freisteht, ungeeignete MCID abzulehnen.

Das Bemühen um eine einheitliche und einfache Regelung ist zu begrüßen. Dennoch halten wir es für sinnvoller, für die Zeit bis zum Vorliegen von akzeptierten Kriterien, nach denen die Qualität einer MCID beurteilt werden kann, bei der bisherigen Vorgehensweise zu bleiben. Nur so bleibt sichergestellt, dass der Bandbreite von Instrumenten und der Vielfalt von therapeutischen Kontexten Rechnung getragen wird und Rückschlüsse auf die Patientenrelevanz aus Ergebnissen komplexer Skalen getroffen werden können.

Literatur

1. Gemeinsamer Bundesausschuss (G-BA). Beschluss des Gemeinsamen Bundesausschusses über die Einleitung eines Stellungnahmeverfahrens zur Änderung der Verfahrensordnung: Änderung der Modulvorlagen in der Anlage II zum 5. Kapitel. 2021.
2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden. Version 6.0 vom 05.11.2020. 2020.
3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden 6.0. 2020.
4. Gemeinsamer Bundesausschuss (G-BA). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über die Einleitung eines Stellungnahmeverfahrens zur Änderung der Verfahrensordnung: Änderung der Modulvorlage in der Anlage II zum 5. Kapitel. 2021.
5. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Magnetresonanztomografiegesteuerte hochfokussierte Ultraschalltherapie zur Behandlung des Uterusmyoms (Addendum zu den Aufträgen E14-04 und E14-05). 2017.
6. Schürmann C. Methodische Aspekte bei der Analyse von Daten zur Lebensqualität in Nutzenbewertungen. IQWiG im Dialog 2021; virtuelle Konferenz 2021.