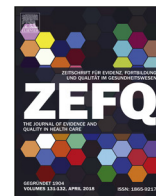




Contents lists available at ScienceDirect

Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ)

journal homepage: <http://www.elsevier.com/locate/zefq>

Evidenz in der Gesundheitsversorgung / Evidence Health Care

Entwicklung von Kriterien für die prospektive Einschätzung des Aktualisierungsbedarfs von Leitlinienempfehlungen: AGIL-Kriterien



Development of criteria for the prospective assessment of the need for updating guideline recommendations: The AGIL criteria

Waldemar Siemens^{a,b,*}, Sonja Mahler^a, Corinna Schaefer^c, Monika Nothacker^d,
Vanessa Piechotta^e, Peggy Prien^c, Sabine Schüller^c, Sabine Schwarz^c, Susanne Blödt^d,
Iris Thielemann^e, Thomas Harder^e, Philipp Kapp^a, Valérie Labonté^{a,b}, Joerg J. Meerpohl^{a,b},
Cordula Braun^{a,b}

^a Institut für Evidenz in der Medizin, Universitätsklinikum Freiburg, Medizinische Fakultät, Albert-Ludwigs-Universität Freiburg, Freiburg, Deutschland^b Cochrane Deutschland, Cochrane Deutschland Stiftung, Freiburg, Deutschland^c Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Berlin, Deutschland^d Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Institut für Medizinisches Wissensmanagement (AWMF-IMWi), Marburg, Deutschland^e Robert Koch-Institut, Berlin, Deutschland

ARTIKEL INFO

Artikel-Historie:

Eingegangen: 21. September 2023

Revision eingegangen: 13. November 2023

Akzeptiert: 21. November 2023

Online gestellt: 18. Januar 2024

Schlüsselwörter:

Leitlinienempfehlungen

Aktualisierung

Priorisierung

Tool

Prospektiv

Survey

Befragung

Checkliste

ZUSAMMENFASSUNG

Hintergrund: Evidenzbasierte Leitlinien- und Impfempfehlungen sollten fortlaufend auf dem aktuellsten Stand sein, um eine angemessene Gesundheitsversorgung zu unterstützen. Gleichzeitig sind die Ressourcen für die Aktualisierung von Leitlinien begrenzt.

Ziel des Projekts war es, eine Kriterienliste für die *prospektive* Einschätzung des Aktualisierungsbedarfs einzelner Leitlinien- oder Impfempfehlungen zu entwickeln, die zum Zeitpunkt der Fertigstellung einer Leitlinie oder Leitlinienaktualisierung anwendbar sind.

Methoden: In diesem Artikel wird die methodische Entwicklung der sogenannten AGIL Kriterien beschrieben (AGIL: Aktualisierungspriorität von Leitlinienempfehlungen; Innovationsfonds-Fördernr.: 01VSF20019). Die AGIL Kriterien wurden in fünf Schritten entwickelt: 1) Erarbeitung einer initialen Kriterienliste innerhalb des Projektteams; 2) Durchführung eines Online-Surveys zur Befragung von Leitlinienexpert*innen zur initialen Kriterienliste; 3) Überarbeitung der Kriterienliste auf der Basis der Ergebnisse des Online-Surveys; 4) Durchführung eines Workshops zur Kriterienliste im Rahmen des EbM-Kongresses 2023; 5) Erstellung von Version 1.0 der AGIL-Kriterien, basierend auf den Ergebnissen des Workshops.

Ergebnisse: Die initiale Liste umfasste die folgenden drei Kriterien: 1) Relevanz der Fragestellung, 2) Verfügbarkeit neuer relevanter Evidenz und 3) Einfluss potenzieller neuer Evidenz.

Die Rücklaufquote des Online-Surveys für vollständig ausgefüllte Fragebögen betrug 31,0 % (N = 195; 630 Leitlinienexpert*innen wurden per E-Mail kontaktiert). Für 90,3 % (n = 176) der Befragten enthielt die Kriterienliste alle für die Einschätzung des Aktualisierungsbedarfs von Leitlinienempfehlungen wesentlichen Aspekte. Mehr als drei Viertel der Befragten bewerteten die Wichtigkeit der drei Kriterien als „sehr wichtig“ oder „wichtig“ (Kriterium 1–3: 75,3 %, 86,1 %, 85,2 %) und – mit Ausnahme von Kriterium 1 – die Verständlichkeit als „sehr verständlich“ oder „verständlich“ (Kriterium 1–3: 58,4 %, 75,9 %, 78,5 %).

Die Ergebnisse des Online-Surveys und des Workshops bestätigten insgesamt die drei Kriterien und ihre jeweils zwei Unterfragen. Die Einarbeitung sämtlichen Feedbacks resultierte in den AGIL-Kriterien (Version 1.0), zusammengefasst: 1) Relevanz der Fragestellung bezüglich a) PICO-Komponenten und b) weiterer Faktoren, z.B. epidemiologische Aspekte; 2) Verfügbarkeit neuer Evidenz a) zum gesundheitsbezogenen Nutzen und Schaden und b) zu anderen Entscheidungsfaktoren, z.B. Umsetzbarkeit, Akzeptanz; 3) Einfluss neuer Evidenz a) auf die Vertrauenswürdigkeit der Evidenz, auf der die Empfehlung basiert, und b) auf die vorliegende Empfehlung, z.B. Empfehlungsstärke.

* Korrespondenzadresse: Dr. Waldemar Siemens, MSc. Institut für Evidenz in der Medizin, Universitätsklinikum Freiburg, Breisacher Str. 86, 79110 Freiburg, Deutschland.
E-Mail: waldemar.siemens@uniklinik-freiburg.de (W. Siemens).

Diskussion: Die moderate Rücklaufquote des Online-Surveys könnte die Repräsentativität eingeschränkt haben. Dennoch bewerten wir die Rücklaufquote in diesem Forschungskontext als zufriedenstellend. Der Einbezug vieler Expert*innen im Rahmen des Online-Surveys und des EbM-Kongress-Workshops stellt eine Stärke des Projekts dar und unterstützt die Qualität der Ergebnisse.

Schlussfolgerungen: Die AGIL-Kriterien bieten eine strukturierte Hilfestellung für die prospektive Einschätzung des Aktualisierungsbedarfs einzelner Leitlinienempfehlungen und anderer evidenzbasierter Empfehlungen. Die Implementierung und Evaluierung der AGIL-Kriterien 1.0 im Rahmen eines Praxistests ist in Planung.

ARTICLE INFO

Article History:

Received: 21 September 2023

Received in revised form: 13 November 2023

Accepted: 21 November 2023

Available online: 18 January 2024

Keywords:

Guideline recommendations

Updating

Prioritization

Tool

Prospective

Survey

Questionnaire

Checklist

ABSTRACT

Background: Evidence-based guideline and vaccination recommendations should continuously be updated to appropriately support health care decisions. However, resources for updating guidelines are often limited.

The aim of this project was to develop a list of criteria for the *prospective* assessment of the need for updating individual guideline or vaccination recommendations, which can be applied from the time a guideline or guideline update is finalised.

Methods: In this article we describe the development of the AGIL criteria (Assessment of Guidelines for Updating Recommendations). The AGIL-criteria were developed by experienced scientists and experts in the field of guideline development in a multi-step process. The five steps included: 1) development of an initial list of criteria by the project team; 2) online survey of guideline experts on the initial version of the criteria list; 3) revision of the criteria list based on the results of the online survey; 4) workshop on the criteria list at the EbM Congress 2023; 5) creation of version 1.0 of the AGIL criteria based on the workshop results.

Results: The initial list included the following three criteria: 1) relevance of the question 2) availability of new relevant evidence, and 3) impact of potentially new evidence.

The response rate of the online survey for fully completed questionnaires was 31.0 % (N = 195; 630 guideline experts were contacted by email). For 90.3 % (n = 176) of the respondents, the criteria list included all essential aspects for assessing the need for updating guideline recommendations. More than three quarters of respondents rated the importance of the three criteria as “very important” or “important” (criteria 1–3: 75.3 %, 86.1 %, 85.2 %) and – with the exception of criterion 1 – comprehensibility as “very comprehensible” or “comprehensible” (criteria 1–3: 58.4 %, 75.9 %, 78.5 %). The results of the online survey and the workshop generally confirmed the three criteria with their two sub-questions. The incorporation of all feedback resulted in the AGIL criteria (version 1.0), recapping: 1) relevance of the question regarding a) PICO components and b) other factors, e. g. epidemiological aspects; 2) availability of new evidence a) on health-related benefits and harms and b) on other decision factors, e. g. feasibility, acceptability; 3) impact of new evidence a) on the certainty of evidence on which the recommendation is based and b) on the present recommendation, e. g. strength of recommendation.

Discussion: The moderate response rate of the online survey may have limited its representativeness. Nevertheless, we consider the response rate to be satisfactory in this research context. The inclusion of many experts in the online survey and the EbM Congress workshop is a strength of the project and supports the quality of the results.

Conclusions: The AGIL criteria provide a structured guidance for the prospective assessment of the need for updating individual guideline recommendations and other evidence-based recommendations. The implementation and evaluation of the AGIL criteria 1.0 in a field test is planned.

Einleitung

Hintergrund

Leitlinien sind nach dem Regelwerk der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) definiert als „systematisch entwickelte Aussagen, die den gegenwärtigen Erkenntnisstand wiedergeben, um die Entscheidungsfindung von Ärzt*innen sowie Angehörige von weiteren Gesundheitsberufen und Patient*innen/Bürger*innen für eine angemessene Versorgung bei spezifischen Gesundheitsproblemen zu unterstützen“ [1]. Um dies zu gewährleisten, müssen evidenzbasierte Leitlinien (z.B. S2e, S3 gemäß AWMF-Stufenklassifikation in Deutschland) und andere evidenzbasierte Empfehlungen, wie z. B. Impfempfehlungen, dem aktuellen Stand des Wissens entsprechen. Das zentrale methodische Instrument zur Identifizierung und Bewertung von Evidenz zur Aktualisierung von Leitlinien sind systematische Übersichtsarbeiten [2].

Das Regelwerk der AWMF fordert, Leitlinien spätestens alle 5 Jahre zu aktualisieren. Für „Living Guidelines“ („Lebende Leitlinien“) sieht das Regelwerk eine jährliche Überprüfung vor [1]. Bei der Aktualisierung einer Leitlinie muss in der Regel eine Priorisierung der zu aktualisierenden Leitlinienempfehlungen stattfinden, um die oftmals limitierten Ressourcen gezielt für die Empfehlungen mit dringendem Aktualisierungsbedarf einzusetzen [2–4]. Die Aktualisierung von Leitlinien ist somit als ein dynamischer Prozess zu verstehen, der mit einer unterschiedlichen Bewertung des Priorisierungsbedarfs der einzelnen Empfehlungen einhergehen kann. Die Priorisierung von Leitlinienempfehlungen im Rahmen von „Living Guideline Recommendations“ (Leitlinienempfehlungen, die bei Erscheinen von neuer Evidenz aktualisiert werden [2]), hat durch die SARS-CoV-2-Pandemie an Bedeutung gewonnen [5–7].

Sanabria et al. publizierten 2020 ein sechs Kriterien umfassendes Tool zur Priorisierung klinischer Fragestellungen im Kontext der Aktualisierung einer Leitlinie, das UpPriority Tool [8]. Als

relevante Kriterien wurden definiert: (1) Auswirkungen veralteter Empfehlungen auf die Sicherheit, (2) Verfügbarkeit neuer relevanter Evidenz, (3) Kontextrelevanz der klinischen Frage, (4) methodische Anwendbarkeit der klinischen Fragestellung, (5) Interesse der Nutzer*innen und (6) Auswirkungen auf den Zugang zur Gesundheitsversorgung. Cheyne et al. stellen in einer aktuellen Arbeit verschiedene Prinzipien zur Identifikation, Selektion und Priorisierung von klinischen Fragen für „Living Guidelines“ dar. Diese Zusammenfassung ist eine Orientierungshilfe, in welcher verschiedene Faktoren beschrieben werden, jedoch kein konkretes Tool zur Bewertung von Empfehlungen [9].

Für die Planung von lebenden Reviewprozessen bzw. die Initiierung von lebenden systematischen Übersichtsarbeiten erscheint die *prospektive* Priorisierung von Leitlinienempfehlungen besonders relevant zu sein. Mit „prospektiv“ ist dabei die Einschätzung des voraussichtlichen Aktualisierungsbedarfs *zum Zeitpunkt der Fertigstellung einer Leitlinie oder Leitlinienaktualisierung* gemeint. Dies unterscheidet sich von der bislang üblicheren Einschätzung des Aktualisierungsbedarfs *zum Zeitpunkt des Beginns der Aktualisierung einer Leitlinie* [8].

Unserer Kenntnis nach gibt es bislang keine Hilfestellung für die *prospektive* Einschätzung des Aktualisierungsbedarfs einzelner Leitlinienempfehlungen. Eine solche Hilfestellung kann die Planung der Aktualisierung von Leitlinien bzw. einzelnen Leitlinienempfehlungen vor dem Hintergrund der verfügbaren Ressourcen erleichtern.

Zielsetzung

Ziel dieses Projekts war es, Kriterien für die *prospektive* Einschätzung des Aktualisierungsbedarfs einer einzelnen Leitlinien- und Impfempfehlung zu entwickeln, das heißt, für die Anwendung zum Zeitpunkt der Fertigstellung einer Leitlinie oder Leitlinienaktualisierung.

Methoden

In diesem Artikel wird die Entwicklung der sog. **Aktualisierungspriorität** von Leitlinienempfehlungen- (AGIL)-Kriterien beschrieben, die im Rahmen des vom Innovationsfonds geförderten DEAL-Projektes erfolgte, Förderkennzeichen 01VSF20019 [10]. Das Akronym „DEAL“ steht für „Dynamische Evidenzaktualisierung für Aktuelle Leitlinienempfehlungen“. Am DEAL-Projekt waren Mitarbeitende des Instituts für Evidenz in der Medizin am Universitätsklinikum Freiburg (IfEM, Konsortialführung) mit Expertise im Erstellen von Systematischen Reviews und Anwendung des GRADE (Grading of Recommendations, Assessment, Development, and Evaluations)-Ansatzes beteiligt. Darüber hinaus waren Mitarbeitende des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ, Konsortialpartner), der Geschäftsstelle der Ständigen Impfkommission (STIKO) des Robert Koch-Instituts (RKI, Konsortialpartner) und des Instituts für Medizinisches Wissensmanagement der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF-IMWi, Kooperationspartner) beteiligt, die die Expertise in der Leitlinienerstellung abdecken.

Methodische Entwicklung

Die folgende Beschreibung der methodischen Entwicklung der AGIL-Kriterien ist in fünf Schritte unterteilt:

1) Entwicklung des ersten Entwurfs der Kriterien durch das Projektteam mittels Literatursuche nach Publikationen zur Priorisierung von Leitlinienempfehlungen bei Leitlinienaktualisierungen und durch die Konsultation von Expert*innen auf dem Gebiet der Leitlinienaktualisierung. Hierzu wurde ein erster Vorschlag von

den beteiligten Mitarbeitenden des IfEM erarbeitet, der im Rahmen mehrerer Online-Konferenzen und Feedback-Schleifen per E-Mail zwischen Januar 2022 und März 2022 gemeinsam von allen Mitgliedern des Projektteams weiterentwickelt und konsentiert wurde. Die initiale Version der Kriterienliste wurde Ende März 2022 erstellt (siehe [Appendix A](#), Tabelle A.1). Diese Version diente als Grundlage für den nachfolgenden Online-Survey. Das UpPriority Tool [8], das sich in Abgrenzung zu den AGIL-Kriterien auf die Priorisierung von Leitlinienempfehlungen *zum Zeitpunkt des Beginns der Leitlinienaktualisierung* bezieht, diente als Orientierung für die Entwicklung der AGIL-Kriterien (Dezember 2021 bis Januar 2022). Die initiale Version der Kriterienliste beinhaltete drei Kriterien: 1) Relevanz der Fragestellung, 2) Verfügbarkeit neuer relevanter Evidenz und 3) Einfluss potenzieller neuer Evidenz. Diese wurden jeweils als Frage mit zwei Unterfragen formuliert und mittels 7-stufiger Likert-Skala bewertet (mit [Appendix A](#) verlinken) (Schritt 1).

2) Mit dem Online-Survey sollte eine möglichst große Zahl von Leitlinienexpert*innen durch das ÄZQ, das AWMF-IMWi und der STIKO in die weitere Entwicklung der AGIL-Kriterien einbezogen werden. Der Online-Survey wurde an Verteiler der vom ÄZQ koordinierten Nationalen VersorgungsLeitlinien, an Leitlinienkoordinierende und Leitlinienbeauftragte von AWMF-Fachgesellschaften sowie an die Mitglieder der STIKO versandt. Der Survey umfasste Items zur Vollständigkeit, Wichtigkeit und Verständlichkeit der Kriterien. Als Leitlinienexpert*innen wurden Personen betrachtet, die aktiv an der Entwicklung oder Aktualisierung einer oder mehrerer Leitlinien bzw. Impfempfehlungen mitgearbeitet hatten.

3) Die Survey-Ergebnisse wurden dazu genutzt, die initiale Kriterienliste weiterzuentwickeln. Sämtliche Änderungen wurden von den Mitgliedern des Projektteams in Online-Konferenzen und per E-Mail konsentiert.

4) Zusätzlich wurde die in Schritt 3 überarbeitete Version der AGIL-Kriterien in einem 90-minütigen Workshop auf der 24. Jahrestagung des Netzwerks Evidenzbasierte Medizin (EbM) vorgestellt und in vier Kleingruppen diskutiert (22.–24.03.2023, Potsdam). Der Workshop stand allen Teilnehmenden der Jahrestagung offen. Angesprochen werden sollten insbesondere Personen mit Erfahrung in der Erstellung von Leitlinien. Als Leitfaden der Gruppendiskussion diente die Einschätzung der Verständlichkeit und Wichtigkeit der AGIL-Kriterien, die chronologisch besprochen wurden, beginnend mit Kriterium 1 (Relevanz der Fragestellung). Abschließend wurden die Ergebnisse in der Gesamtgruppe durch Präsentationen der einzelnen Gruppen zusammengeführt und dokumentiert [11].

5) Die im Rahmen des Workshops eingeholten Rückmeldungen wurden zusammengefasst und Änderungen innerhalb des Projektteams in Online-Konferenzen und per E-Mail konsentiert. Dies resultierte letztlich in den AGIL-Kriterien (Version 1.0).

Online-Survey

Der Online-Survey gliederte sich in vier Abschnitte (siehe [Appendix A](#), Zusatzinformation A.1):

1) Vier allgemeine Fragen zur Person (z. B. Berufsgruppe).
2) Drei Fragen zur Kriterienliste insgesamt im Sinne eines Gesamteindrucks und der Vollständigkeit.

3.1) Sechs geschlossene Fragen zur Einschätzung der Wichtigkeit und Verständlichkeit der drei Kriterien mittels einer 7-stufigen Likert-Skala.

3.2) Je Kriterium eine offene Frage zu ergänzenden Anmerkungen und Kommentaren zur Wichtigkeit und Verständlichkeit der drei Kriterien (siehe [Appendix A](#), Zusatzinformation A.1).

4. Zudem gab es die Möglichkeit, am Ende des Fragebogens ergänzende Anmerkungen und Gedanken zur Kriterienliste in einer offenen Frage zu äußern.

Für den Survey wurde ein Ethikantrag bei der Ethikkommission des Universitätsklinikums Freiburg gestellt (Antrags-Nr. 22-1150). Das positive Votum wurde am 19.05.2022 erteilt.

Die Befragung wurde mit SoSci-Survey (<https://www.soscisurvey.de>) im Zeitraum vom 05.10.2022 bis zum 12.11.2022 durchgeführt. Die Auswertung wurde im Januar 2023 abgeschlossen. Die Teilnahme am Online-Survey war freiwillig und anonym. Der Versand des Teilnahme-Links erfolgte über die drei Projektpartner AWMF-IMWi, ÄZQ und RKI, indem ein Einladungslink in SoSci-Survey erstellt und per E-Mail über die Verteiler der Partner*innen verschickt wurde. Nach zwei Wochen und nochmals nach vier Wochen wurden an alle Adressat*innen Erinnerungsmails gesendet, um den Rücklauf der beantworteten Fragebögen zu erhöhen. Für die Auswertung wurden nur Fragebögen mit vollständigem quantitativen Teil herangezogen.

Die Ergebnisse des Online-Surveys werden unter Berücksichtigung der Kriterien der Consensus-Based Checklist for Reporting of Survey Studies (CROSS) berichtet [12].

Datenanalyse

Die erhobenen Daten wurden mit Microsoft Excel 2016 ausgewertet. Die quantitativen Ergebnisse des Online-Surveys (Antworten auf die geschlossenen Fragen) wurden mittels deskriptiver Statistik ausgewertet (absolute und relative Häufigkeiten) und in Tabellen sowie Diagrammen dargestellt.

Die qualitativen Ergebnisse der Umfrage (Antworten auf acht offene Fragen) (Appendix A, Zusatzinformation A.1) wurden gemäß der inhaltlich strukturierenden Inhaltsanalyse nach Kuckartz ausgewertet [13]. Der grundsätzliche Ablauf beinhaltet sieben Schritte, die von der Identifikation wichtiger Textstellen, dem Entwickeln, Codieren und Zusammenstellen von Hauptkategorien bis zum induktiven Bestimmen von Subkategorien, der Ausdifferenzierung des Kategoriensystems und der abschließenden Ergebnisdarstellung reichen. Eine Person (SM) fasste die Daten zusammen und bildete in Abstimmung mit zwei weiteren Wissenschaftler*innen des Projektteams (WS, CB) die Ober- und Unterkategorien. Die Oberkategorien der qualitativen Analyse orientierten sich an der Darstellung der einzelnen Kriterien im Online-Survey (Appendix A, Tabelle A.1) und wurden deshalb in drei Teile unterteilt: Fragestellung, Frage a) und Frage b). Zusätzlich wurden Kriterien-übergreifende Kategorien gebildet, wie z.B. Anmerkungen zur Sprache und Gewichtung der einzelnen Kriterien. Nach Zuordnung der gegebenen Antworten in die Oberkategorien wurden Unterkategorien basierend auf den inhaltlichen Schwerpunkten der vorliegenden Antworten gebildet.

Ergebnisse

Online-Survey

Quantitative Ergebnisse

Für die Befragung (Schritt 2) wurden 630 Leitlinienexpert*innen kontaktiert (ÄZQ: n = 331, AWMF-IMWi: n = 277, STIKO: n = 22). Insgesamt wurde das Ausfüllen von 232 Online-Surveys begonnen, 195 Fragebögen wurden bezüglich der quantitativen Fragen vollständig ausgefüllt und ausgewertet (=31,0% Rücklaufquote), 37 Leitlinienexpert*innen brachen den Fragebogen vorzeitig ab und wurden nicht in die Auswertung eingeschlossen.

Im ersten Teil des Online-Surveys wurden allgemeine Eigenschaften zur Person abgefragt (siehe Tabelle 1). Mit 68,2% (n = 133) war der größte Teil der Teilnehmenden Ärzt*innen und 86,7% (n = 169) der Teilnehmenden gaben an, Mandatsträger*in einer professionellen Fachgesellschaft/Organisation zu sein. Die Mehrheit (70,8%, n = 138) der Befragten hatte bereits aktiv an

Tabelle 1

Stichprobenbeschreibung (N = 195).

Fragen und Antwortmöglichkeiten	n	%
Welcher der folgenden Gruppen würden Sie sich primär zuordnen?		
Arzt/Ärztin	133	68,2
Mitglied eines Heilberufs (z.B. Pharmazeut*in, Psychotherapeut*in)	12	6,2
Mitglied eines Gesundheitsfachberufs (Physiotherapeut*in, Logopäd*in)	19	9,7
Wissenschaftler*in/Methodiker*in	14	7,2
Patient*innenvertreter*in	10	5,1
Sonstige [#]	7	3,6
In welcher/n Rolle/n haben Sie bislang an der Entwicklung bzw. Aktualisierung von Leitlinien und/oder Impfpfehlungen mitgewirkt? (Mehrfachantworten) [§]		
Mandatsträger*in einer professionellen Fachgesellschaft/Organisation	169	86,7
Mitarbeiter*in einer Leitlinien- oder impfpfehlungerstellenden Institution	57	29,2
Externe Leitlinienberater*in	17	8,7
Mitarbeiter*in einer Evidenz-aufbereitenden Institution	16	8,2
Sonstige [§]	17	8,7
An wie vielen Leitlinien- oder Impfpfehlungs-Projekten haben Sie bislang aktiv mitgewirkt?		
1–2	57	29,2
3–5	78	40,0
6–10	34	17,4
>11	26	13,3
Wenn Sie bereits an Leitlinien- oder Impfpfehlungs-Projekten mitgewirkt haben:		
Haben Sie bereits aktiv an Aktualisierungs- oder Überarbeitungsprozessen mitgewirkt?		
Ja	173	88,7
Nein	22	11,3

[#] mehrere der oben genannten gleichzeitig (n = 1), Sozialwissenschaft (n = 1), Verband (n = 1), Leitlinienbeauftragte (n = 1), Gesundheitswissenschaftler*in (n = 1), Pflege (n = 1), Gesundheitsfachberuf Pflegefachperson (n = 1)

[§] Anzahl der gegebenen Antworten: 1 (n = 132, 67,7%), 2 (n = 50, 25,6%), 3 (n = 8, 4,1%), 4 (n = 5, 2,6%); gesamt N = 276

[§] Leitlinienkoordinator*in (n = 6), Patient*innenvertreter*in (n = 4), Autor*in (n = 1), Federführende*r Autor*in (n = 1), Evidenzaufbereitung für Lehre (n = 1), Sportmedizin (n = 1), Leitlinienbeauftragte (n = 1), Steuerungsgruppe (n = 2).

drei oder mehr Leitlinien oder Impfpfehlungen mitgewirkt. Ein Großteil der Befragten (88,7%, n = 173) hatte bereits an Aktualisierungs- oder Überarbeitungsprozessen mitgearbeitet.

Der zweite Teil des Online-Surveys bezog sich auf die Kriterienliste im Ganzen. 90,3% (n = 176) der Befragten stimmten der Aussage zu, dass die Kriterienliste alle wesentlichen Aspekte enthält, die für die Einschätzung des zukünftigen Aktualisierungsbedarfs einer einzelnen Leitlinien- oder Impfpfehlung wichtig sind. Auf einer 7-stufigen Likert-Skala gab ein Großteil der Expert*innen (78,5%) an, dass die Kriterienliste „weder zu wenige noch zu viele“ Kriterien enthält. 8,2% der Befragten kamen zur Einschätzung, dass es „eher zu wenige“ und 8,7%, dass es „eher zu viele“ Kriterien sind. Die übrigen Antwortkategorien („viel zu viele“, „zu viele“, „zu wenige“, „viel zu wenige“) beliefen sich auf weniger als 3% pro Kategorie (siehe Abbildung 1).

Mehr als drei Viertel der Befragten bewerteten die Wichtigkeit der drei Kriterien als „sehr wichtig“ oder „wichtig“ (Kriterium 1–3: 75,3%, 86,1%, 85,2%) und – mit Ausnahme von Kriterium 1 – die Verständlichkeit als „sehr verständlich“ oder „verständlich“ (Kriterium 1–3: 58,4%, 75,9%, 78,5%) (Abbildung 2).

Qualitative Ergebnisse

Die qualitative Analyse basierte auf den Antworten der Expert*innen auf offene Fragen zur Wichtigkeit beziehungsweise zur Verständlichkeit der gesamten Kriterienliste (Teil 2 des Online-Surveys) und der einzelnen Kriterien (Teil 3 des Online-Surveys, Appendix A, Zusatzinformation A.1). Die qualitativen Fragen wurden am häufigsten (n = 33) für „3.6 Möchten Sie etwas zu Ihrer

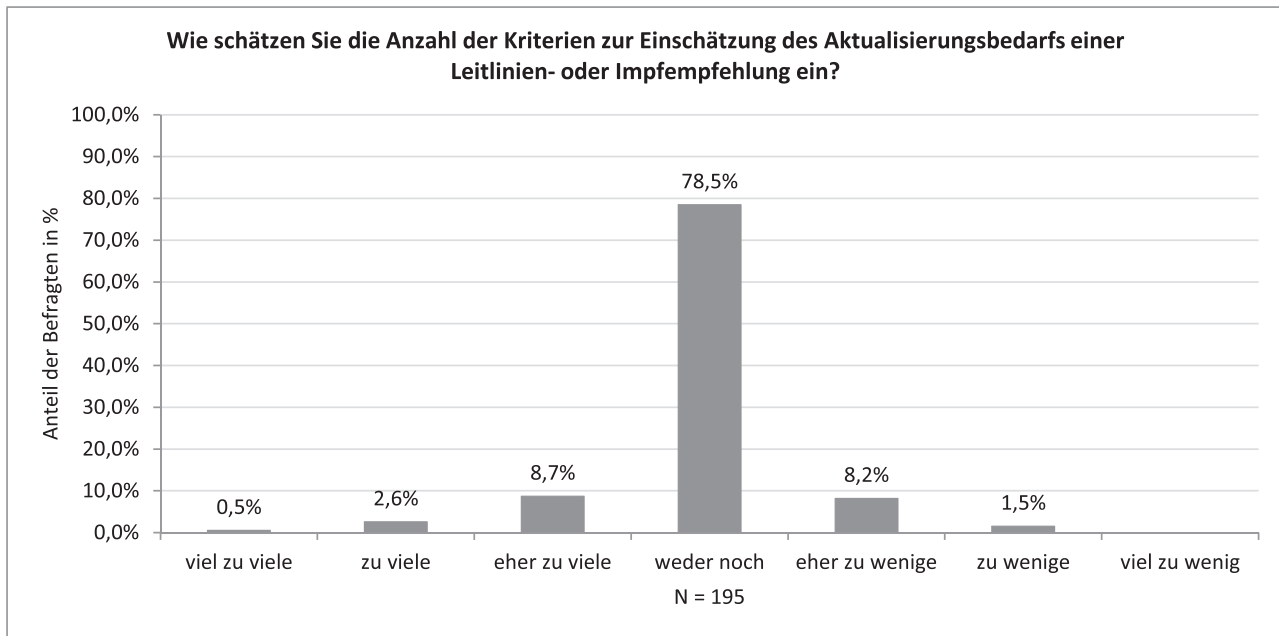


Abbildung 1. Einschätzung der Anzahl der Kriterien.

Einschätzung der Wichtigkeit des Kriteriums 2 anmerken?“ und am wenigsten häufig ($n = 11$) für Item „3.12 Würden Sie das Kriterium 3 anders formulieren, damit es verständlicher ist?“ ausgefüllt.

Eine zusammenfassende Darstellung der Ober- und Unterkategorien der drei Kriterien ist Abbildung A.1 im [Appendix A](#) zu entnehmen. Zum inhaltlichen Verständnis der Oberkategorien und Unterkategorien wird auf [Appendix A](#), Tabelle A.1 verwiesen, in dem jede Unterkategorie durch ein exemplarisches Zitat aus der Befragung gestützt ist.

In Bezug auf Kriterium 1 („Relevanz der Fragestellung“), das die zukünftige Relevanz der Fragestellung für Patient*innen, die Zielpopulation und relevante Akteur*innen (z. B. Gesundheitsfachpersonen) betrifft, wurde angemerkt, dass die Begriffe „Zielpopulation“ und „Patient*innen“ zu Missverständnissen führen könnten (TN Nr. 612). Daher wurde der allgemeinere Begriff „Zielpopulation“ für die AGIL-Kriterien (Version 1.0) gewählt, der sowohl Fragestellungen zur Therapie als auch zur Prävention umfasst ([Tabelle 2](#)). Darüber hinaus wurde angemerkt, dass die Bewertung des Begriffs „Relevanz“/„relevant“ nicht unproblematisch sei, da es für diesen keine allgemein gültige Definition gibt. So wurde einerseits betont, dass die Fragestellung immer dieselbe Relevanz habe (TN Nr. 567), andererseits wurde die Notwendigkeit einer Aktualisierung gerade mit der Veränderung der Relevanz der Fragestellung in Verbindung gebracht (TN Nr. 378). In Bezug auf die Verwendung des Begriffs „Kontextfaktoren“ in Frage „b) in Bezug auf Kontextfaktoren (z. B. Krankheitslast, Veränderungen in der klinischen Praxis oder neue Versorgungsmöglichkeiten)“ von Kriterium 1 wurde von einer Person der Wunsch nach einer präziseren Definition des Begriffs „Kontextfaktoren“ und nach Erweiterung um „Krankheitshäufigkeit sowie die Häufigkeit schwerer Verläufe“ geäußert (TN Nr. 341). Zudem wurde von einer Person die Ergänzung von „epidemiologischen Aspekten“ als Kontextfaktor vorgeschlagen (TN Nr. 359). Insgesamt führten die Anmerkungen zur Änderung in den breiteren Begriff „Faktoren“

statt „Kontextfaktoren“ und zum Hinzufügen des Begriffs „epidemiologische Aspekte“ als Beispiel der Unterfrage 1 b) ([Tabelle 2](#)).

Zum zweiten Kriterium („Verfügbarkeit neuer Evidenz“) wurde angemerkt, dass der Begriff „Relevanz“ in der Überschrift „Verfügbarkeit neuer relevanter Evidenz“ ähnliche Probleme wie bei Kriterium 1 aufweist (TN Nr. 487). Die Definition von „relevanter Evidenz“ kann unterschiedlich interpretiert werden und muss innerhalb einer Leitliniengruppe individuell definiert werden. Der Begriff „wahrscheinlich“ wurde von einer Person als „komplex“ empfunden (TN Nr. 450), spiegelt jedoch den prospektiven Ansatz der AGIL-Kriterien bezüglich der Aktualisierung der Leitlinienempfehlungen zum Zeitpunkt der Fertigstellung oder Aktualisierung einer Leitlinie oder Leitlinienempfehlung wider. Es wurde deshalb von der DEAL-Projektgruppe entschieden, den Begriff „wahrscheinlich“ als Bestandteil des Einleitungssatzes für alle drei Kriterien beizubehalten („Wie wahrscheinlich ist...“). Der Begriff „zeitnah“ wurde als zu allgemein wahrgenommen (TN Nr. 450) und daher mit einer erklärenden Fußnote versehen ([Tabelle 2](#)).

Für Kriterium 3 („Einfluss neuer Evidenz“) wurde angemerkt, dass der Begriff „Vertrauenswürdigkeit“ in „weitere Gültigkeit oder Relevanz“ geändert werden könne (TN Nr. 650). Da „Vertrauenswürdigkeit“ jedoch ein feststehender Begriff gemäß dem GRADE-Ansatz ist, wurde eine Fußnote zur Erklärung hinzugefügt. Der Begriff „potentiell“ wurde aus der Bezeichnung von Kriterium 3 entfernt, da dieser irritierte. Zudem wurde der Begriff „Subgruppen“ mit einer erklärenden Fußnote versehen ([Tabelle 2](#)).

In Bezug auf die übergreifenden Kategorien äußerten die Befragten den Wunsch, die Formulierung der Kriterien einfacher zu gestalten und auf Klammern und komplexe Satzstrukturen zu verzichten (TN Nr. 361, TN Nr. 425, Oberkategorie „Sprache“, [Appendix A](#), Tabelle A.1), was bestmöglich berücksichtigt wurde ([Tabelle 2](#)). Einige Anmerkungen waren generell hilfreich, überstiegen jedoch den Rahmen der AGIL-Kriterien und wurden daher bewusst nicht implementiert, z. B. der Einfluss gesetzlicher

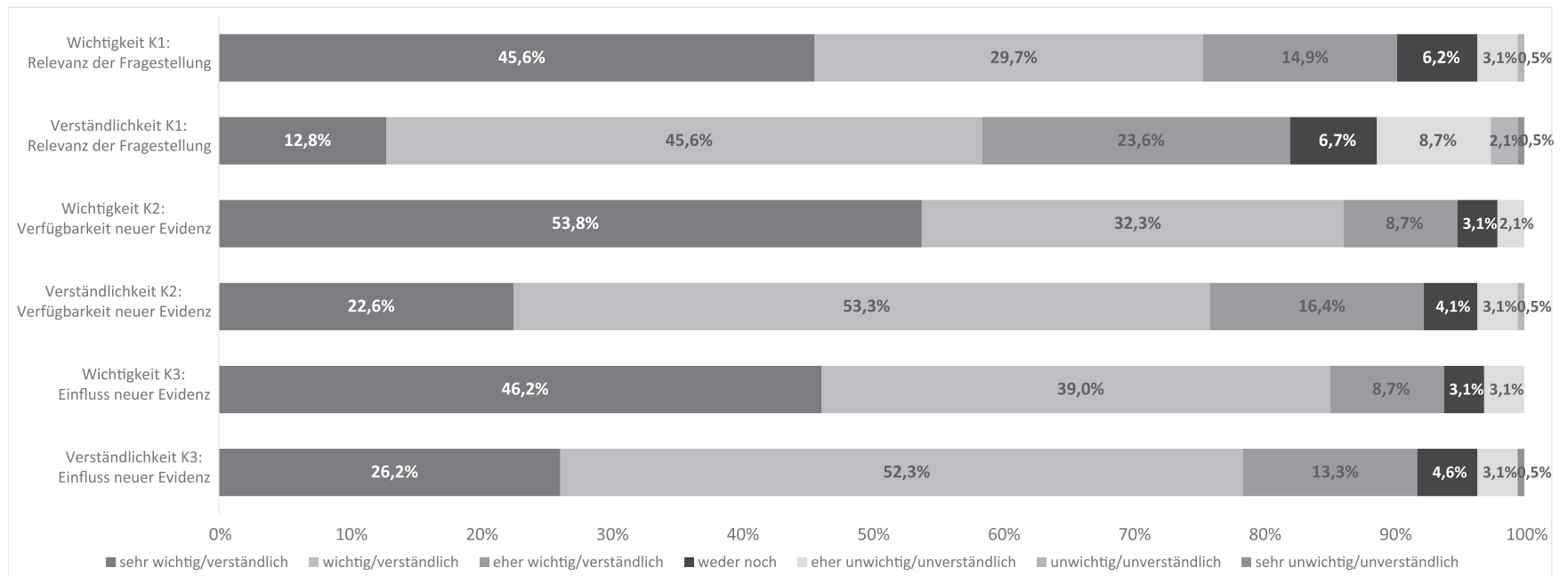


Abbildung 2. Wichtigkeit und Verständlichkeit der AGIL-Kriterien (Version 1.0), N = 195. Bewertungen in Graustufen und Legende einheitlich von links beginnend; K: Kriterium. Kriterium 1 - Relevanz der Fragestellung: Wie wahrscheinlich ist es, dass die Fragestellung für die Zielpopulation und/oder relevante Akteur*innen (z.B. Gesundheitsfachpersonen, Entscheidungsträger*innen) weiterhin bedeutsam sein wird: a) bezogen auf die PICO-Komponenten (Population, Intervention, Vergleich, Endpunkte)? b) bezogen auf weitere Faktoren (z.B. epidemiologische Aspekte, Veränderungen des Gesundheitssystems)? Kriterium 2 - Verfügbarkeit neuer Evidenz: Wie wahrscheinlich ist es, dass neue relevante Evidenz zeitnah verfügbar sein wird: a) zum gesundheitsbezogenen Nutzen und Schaden (ggf. auch bezüglich neuer Subgruppen)? b) zu anderen Entscheidungsfaktoren (z.B. Wertvorstellungen und Präferenzen, Ressourcenverbrauch und Kosten, gesundheitliche Chancengleichheit, Umsetzbarkeit, Akzeptanz)? Kriterium 3 - Einfluss neuer Evidenz: Wie wahrscheinlich ist es, dass neue Evidenz einen bedeutsamen Einfluss haben wird: a) auf die Vertrauenswürdigkeit der Evidenz, auf der die Empfehlung basiert? b) auf die vorliegende Empfehlung (z.B. Änderung der Empfehlungsstärke, Überlegungen zu Subgruppen)?

Tabelle 2

AGIL-Kriterien (Version 1.0) für die prospektive Einschätzung des Aktualisierungsbedarfs einer einzelnen Leitlinien- und Impfpfempfehlung.

Kriterium 1 – Relevanz der Fragestellung						
Wie wahrscheinlich ist es, dass die Fragestellung für die Zielpopulation und/oder relevante Akteur*innen (z.B. Gesundheitsfachpersonen, Entscheidungsträger*innen) weiterhin bedeutsam sein wird						
a) bezogen auf die PICO-Komponenten (Population, Intervention, Vergleich, Endpunkte)?						
sehr unwahrscheinlich	unwahrscheinlich	eher unwahrscheinlich	weder noch	eher wahrscheinlich	wahrscheinlich	sehr wahrscheinlich
0	0	0	0	0	0	0
b) bezogen auf weitere Faktoren (z.B. epidemiologische Aspekte, Veränderungen des Gesundheitssystems)?						
sehr unwahrscheinlich	unwahrscheinlich	eher unwahrscheinlich	weder noch	eher wahrscheinlich	wahrscheinlich	sehr wahrscheinlich
0	0	0	0	0	0	0
Kriterium 2 – Verfügbarkeit neuer Evidenz						
Wie wahrscheinlich ist es, dass neue relevante Evidenz zeitnah ¹ verfügbar sein wird						
a) zum gesundheitsbezogenen Nutzen und Schaden (ggf. auch bezüglich neuer Subgruppen)?						
sehr unwahrscheinlich	unwahrscheinlich	eher unwahrscheinlich	weder noch	eher wahrscheinlich	wahrscheinlich	sehr wahrscheinlich
0	0	0	0	0	0	0
b) zu anderen Entscheidungsfaktoren (z.B. Wertvorstellungen und Präferenzen, Ressourcenverbrauch und Kosten, gesundheitliche Chancengleichheit, Umsetzbarkeit, Akzeptanz)?						
sehr unwahrscheinlich	unwahrscheinlich	eher unwahrscheinlich	weder noch	eher wahrscheinlich	wahrscheinlich	sehr wahrscheinlich
0	0	0	0	0	0	0
Kriterium 3 – Einfluss neuer Evidenz						
Wie wahrscheinlich ist es, dass neue Evidenz einen bedeutsamen Einfluss haben wird						
a) auf die Vertrauenswürdigkeit der Evidenz ² , auf der die Empfehlung basiert?						
sehr unwahrscheinlich	unwahrscheinlich	eher unwahrscheinlich	weder noch	eher wahrscheinlich	wahrscheinlich	sehr wahrscheinlich
0	0	0	0	0	0	0
b) auf die vorliegende Empfehlung (z.B. Änderung der Empfehlungsstärke, Überlegungen zu Subgruppen ³)?						
sehr unwahrscheinlich	unwahrscheinlich	eher unwahrscheinlich	weder noch	eher wahrscheinlich	wahrscheinlich	sehr wahrscheinlich
0	0	0	0	0	0	0

¹ Zeitnah: die Definition von „zeitnah“ sollte vor Nutzung der Liste dem jeweiligen Kontext entsprechend festgelegt werden.

² Eine häufig angewandte Einschätzung der „Vertrauenswürdigkeit der Evidenz“ ist beispielsweise die nach dem Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) Ansatz: hoch, moderat, niedrig, sehr niedrig.

³ Z.B. separate Empfehlungen für Subgruppen bei neuer, relevanter Evidenz.

Vorgaben auf Handlungsempfehlungen (TN Nr. 313) oder Gesichtspunkte der Leistungserstattung (TN Nr. 349). Schließlich diente ein Hinweis von TN Nr. 450 zur Bewertung der Unterfragen mit „sehr wahrscheinlich“ bis „sehr unwahrscheinlich“ als wichtige Anregung, eine 7-stufige Likert-Skala für die Bewertung einer jeden Unterfrage a) und b) zu integrieren (Tabelle 2).

AGIL-Kriterien: Version 1.0

Für die Überarbeitung der Kriterienliste basierend auf den Ergebnissen des Online-Surveys erfolgten Online-Diskussionen im Projektteam und ein Austausch per E-Mail (Schritt 3). Im Rahmen des Workshops auf dem EbM-Kongress 2023 mit rund 30 Teilnehmenden mit Vorerfahrungen in der Leitlinienarbeit wurde die aus Schritt 3 resultierende Kriterienliste in Kleingruppen und nachfolgend in der Gesamtgruppe diskutiert (Schritt 4). Das Feedback der Workshopteilnehmenden wurde systematisch dokumentiert und berücksichtigt [11] (Schritt 5) und führte schließlich zu Version 1.0 der AGIL-Kriterien.

Die Ergebnisse des Online-Surveys (Schritt 3) und des Workshops (Schritt 4) bestätigten die drei Kriterien mitsamt ihren je zwei Unterfragen a) und b). Die wesentlichen Änderungen zwischen der initialen Kriterienliste (Schritt 1) in Tabelle A1, die noch keine Möglichkeiten zur quantitativen Bewertung der Kriterien beinhaltete, und Version 1.0 der AGIL-Kriterien in Tabelle 2, waren das Hinzufügen einer 7-stufigen Likert-Skala zur Einschätzung der Wahrscheinlichkeit („sehr unwahrscheinlich“ bis „sehr wahrscheinlich“) bezüglich der Unterfragen a) und b) jedes Kriteriums. Zudem wurde die Formulierung konkretisiert und es wurden Fußnoten zur Erklärung hinzugefügt. Der Begleittext (Tabelle 3) setzt die AGIL-Kriterien in den Kontext der Anwendung, indem Informationen und eine Abgrenzung zur potenziellen Anwendung einer Gewichtung der Items und/oder der Bildung eines Gesamt-Scores gegeben werden. Der Begleittext in Tabelle 3 verdeutlicht, dass die Bewertung der Kriterien und die daraus folgende Priorisierung individuell für das jeweilige Leitlinien-/Empfehlungsprojekt zu handhaben ist. Die Bildung eines Gesamt-Scores und eines festen Cut-Off-Werts werden nicht empfohlen, weil die Anzahl der zu bearbeitenden Empfehlungen u. a. von den zur Verfügung stehenden Ressourcen abhängt. Aus denselben Gründen werden auch keine Vorgaben zur Gewichtung der einzelnen Kriterien

gemacht. Ob die praktische Anwendung der AGIL-Kriterien mit der Bildung eines Gesamt-Scores oder der Gewichtung von Items stark verbessert werden, wird Teil der geplanten Testung sein.

Diskussion

Die vorliegende Arbeit beschreibt die mehrstufige methodische Entwicklung der AGIL-Kriterien für die prospektive Einschätzung des Aktualisierungsbedarfs einzelner Leitlinien- und Impfeempfehlungen zum Zeitpunkt der Fertigstellung einer Leitlinie oder Leitlinienaktualisierung.

Die Auswertung der quantitativen Daten ergab, dass die meisten der Leitlinienexpert*innen die AGIL-Kriterien als wichtig und insgesamt verständlich einschätzten. Aus den qualitativen Daten des Online-Surveys, dem Workshop bei der Jahrestagung des Netzwerks Evidenzbasierte Medizin [11] und dem Konsensprozess des Projektteams (IfEM, ÄZQ, AWMF-IMWi und RKI) ergaben sich die folgenden wesentlichen Diskussionspunkte, die teilweise zu Änderungen führten: die Anzahl und Reihenfolge der Kriterien und ihrer Unterfragen, die Gestaltung der quantitativen Bewertung der Unterfragen für jedes der drei Kriterien (z.B. 5-stufige vs. 7-stufige Likert-Skala), die Ergänzung von Fußnoten mit Erklärungen, das Verfassen eines Begleittextes mit Hinweisen zur Auswertung, insbesondere zur Gewichtung der Items und/oder der Bildung eines Gesamt-Scores. Im Gegensatz zu Sanabria et al. [8], die für das UpPriority Tool die Berechnung eines Gesamt-Scores vorschlagen, raten wir diesbezüglich zur Vorsicht und zur separaten Betrachtung der Ergebnisse für jedes Kriterium in Übereinstimmung mit anderen Tools zur Bewertung von Evidenz [14–16]. Hintergrund ist, dass die Berechnung eines Gesamt-Scores entweder eine äquivalente Gewichtung aller Fragen voraussetzt oder aber eine – nicht empirisch gesicherte – unterschiedliche Festlegung von Gewichten erfordern würde. Dies scheint aktuell nicht angebracht.

Forschungskontext

Als Orientierung für die Erstellung der initialen Version der AGIL-Kriterien wurde u. a. das UpPriority Tool berücksichtigt [8]. Das UpPriority Tool wurde als pragmatisch und einfach zu verwendendes Tool entwickelt, um die in einer Leitlinie enthaltenen

Tabelle 3

Informationen und Hinweise zur Nutzung der AGIL-Kriterien.

Wozu dient die Kriterienliste?

Die vorliegende Kriterienliste bietet eine Hilfestellung für die vorausschauende Einschätzung des künftigen Aktualisierungsbedarfs einzelner Empfehlungen evidenzbasierter Leitlinien sowie anderer evidenzbasierter Entscheidungshilfen wie z. B. Impfeempfehlungen.

Die Kriterienliste wurde für die Anwendung zum Zeitpunkt der initialen Fertigstellung einer Leitlinie oder der Aktualisierung einer Leitlinie konzipiert. Die vorausschauende Einschätzung des Aktualisierungsbedarfs mehrerer, ggf. sämtlicher, Empfehlungen einer Leitlinie kann für eine ressourcenangepasste Priorisierung der Aktualisierung einzelner Empfehlungen genutzt werden.

Was ist bei der Nutzung der Kriterienliste zu beachten?

Die Kriterienliste wurde primär für die Nutzung von Empfehlungen zu therapeutischen Interventionen entwickelt, ist aber mit geringfügigen Anpassungen auch für Empfehlungen z. B. zu diagnostischen Fragestellungen nutzbar.

Die Bewertung der einzelnen Items erfolgt mit einer 7-stufigen Likert-Skala und ermöglicht die Zuordnung eines numerischen Wertes zwischen 0 (für „sehr unwahrscheinlich“) und 6 („sehr wahrscheinlich“) für jedes Item. Die Auswertung durch Bildung eines Gesamt-Scores oder Gewichtung der Items ist nicht primär vorgesehen, kann aber bei Bedarf und mit Bedacht durchgeführt werden, sofern dies die Anwendbarkeit der Kriterienliste verbessert.

Wer hat die Kriterienliste entwickelt?

Die Kriterienliste wurde im Rahmen eines vom Innovationsfonds geförderten Projektes („DEAL“¹ 01VSF20019) von Mitarbeitenden des Instituts für Evidenz in der Medizin (IfEM²), des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ³), des Instituts für Medizinisches Wissensmanagements der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF-IMWi⁴), und von Mitarbeitenden der Geschäftsstelle der Ständigen Impfkommission (STIKO⁵) beim Robert Koch-Institut entwickelt. Mittels eines Online-Surveys wurden die Mitglieder von Leitliniengruppen und der STIKO in den Prozess der Erstellung der Kriterienliste einbezogen.

¹ <https://innovationsfonds.g-ba.de/projekte/versorgungsforschung/deal-dynamische-evidenzaktualisierung-fuer-aktuelle-leitlinienempfehlungen.399>

² <https://www.uniklinik-freiburg.de/institut-fuer-evidenz-in-der-medicin.html>

³ <https://www.aezq.de/>

⁴ <https://www.awmf.org/die-awmf/imwi>

⁵ https://www.rki.de/DE/Content/Institut/institut_node.html

klinischen Fragen bzw. Empfehlungen zum Zeitpunkt der Aktualisierung zu priorisieren [8], während die AGIL-Kriterien zur prospektiven Einschätzung des Aktualisierungsbedarfs der Empfehlungen einer gerade abgeschlossenen Leitlinie oder Leitlinienaktualisierung entwickelt wurden. Insbesondere im Kontext der Planung von lebenden Reviewprozessen bzw. der Initiierung von lebenden systematischen Übersichtsarbeiten erscheint die prospektive Priorisierung von großer Relevanz. In Übereinstimmung mit Sanabria et al. haben wir ebenfalls die 7-stufige Likert-Skala gewählt, um die Trennschärfe der Kriterien zu erhöhen.

Im Kontext von Living Guidelines haben Cheyne et al. darüber hinaus Leitprinzipien für die Auswahl von Leitlinienempfehlungen vorgeschlagen (z. B. anhand hoher Priorität der Fragestellung und niedriger Vertrauenswürdigkeit der Evidenz) [9]. Weitere Leitprinzipien für die Priorisierung, Überarbeitung oder Entfernung von Leitlinienempfehlungen aus dem „living mode“ werden ausführlich beschrieben und reichen von der Betrachtung von Ressourcen bis zur stetigen Reflexion der Priorität (z. B. Überarbeitung und Kommunikation von prioritären Fragestellungen sowie deren Depriorisierung) [9]. Im Gegensatz zu den AGIL-Kriterien nennen die Autor*innen jedoch keine expliziten Kriterien im Sinne eines Tools, um die Fragen nach Selektion, Priorisierung, Überarbeitung oder Entfernung von Leitlinienempfehlungen aus dem „living mode“ zu beantworten.

Eine weitere 2023 publizierte Arbeit von Fraile Navarro et al. [17] ging der Frage nach, wann neue Evidenz einzubeziehen ist und was Auslöser für die Synthese, Bewertung und Einbindung in Leitlinienempfehlungen sein könnten. Ein ressourcen-intensiver Weg, der jedoch die schnellste Integration neuer Evidenz in den Evidenzkörper darstellt, ist die „kontinuierliche Integration“ neuer Evidenz sobald neue Evidenz zugänglich ist. Dieser Ansatz ist angebracht für Leitlinienempfehlungen der höchsten Priorität, z.B. im Falle sehr großer gesellschaftlicher Relevanz einer Fragestellung. Ein alternativer Ansatz ist die Trigger-basierte („trigger-based“) Integration von Evidenz. Dem Trigger-basierten Ansatz zufolge bestimmen vordefinierte und transparente Kriterien („Trigger“) eine Aktualisierung der Evidenzsynthese, z. B. a) Studien mit einer bestimmten Anzahl an Teilnehmenden und/oder einer hohen Präzision, b) Studien mit Änderungspotential für die Vertrauenswürdigkeit der Evidenz nach dem GRADE-Ansatz, c) Studien mit Erreichen einer minimal klinisch relevanten Differenz oder d) Studien, die aufgrund von Expert*innenmeinungen bzw. eines Expert*innenpanels berücksichtigt werden sollten [17]. Die Autor*innen nennen keine expliziten Items im Sinne eines Tools, sodass die beschriebenen Auslöser-basierten Punkte nicht unmittelbar bewertet und quantifiziert werden können.

Die beschriebenen Ansätze weisen Überschneidungen auf, z. B. im Fokus auf die Relevanz der Leitlinienempfehlung, der Berücksichtigung neuer verfügbarer Evidenz und der Priorisierung von Empfehlungen [8,17,18]. Der Mehrwert der AGIL-Kriterien besteht darin, dass sie durch die geringe Anzahl an Kriterien und der 7-stufigen Likert-Skala kompakt und schnell anwendbar sind. Zudem richtet sich der Blick bei Anwendung der AGIL-Kriterien in die Zukunft und fokussiert sich auf die prospektive Einschätzung („Wahrscheinlichkeit“) des zukünftigen Aktualisierungsbedarfs einzelner Leitlinien- und Impfehlungen.

Stärken und Limitationen

Zu den Stärken dieser Arbeit zählt das systematische, mehrstufige Vorgehen bei der Entwicklung der AGIL-Kriterien. Hierbei ist insbesondere der umfangreiche Online-Survey mit 195 vollständigen Datensätzen von Leitlinienexpert*innen des AWMF-IMWi, der ÄZQ und der STIKO zu nennen. Die Daten wurden sowohl quantitativ mittels deskriptiver Statistik als auch qualitativ mit der inhaltlich strukturierenden Inhaltsanalyse nach Kuckartz

ausgewertet [13]. Dies erlaubte eine vielschichtige und tiefgründige Betrachtung der Ergebnisse. Eine weitere Stärke ist die Vorstellung der Kriterienliste in einem Workshop im Rahmen der 24. Jahrestagung des Netzwerks Evidenzbasierte Medizin. Während des Workshops gaben ca. 30 Teilnehmende ihr Feedback zu den AGIL-Kriterien. Jegliche Änderungen der AGIL-Kriterien wurden von allen Mitgliedern des Projektteams (IfEM, ÄZQ, AWMF-IMWi, RKI) in Online-Konferenzen und per E-Mail konsentiert, was eine weitere Stärke des Projekts darstellt.

Die Stichprobe des vorgestellten Online-Surveys basierte auf den E-Mail-Verteilern von ÄZQ, AWMF-IMWi und RKI, wobei die ersteren beiden mit je ca. 300 teilnehmenden Leitlinienexpert*innen die Stichprobe klar dominierten. Der Anteil der Impfexpert*innen der STIKO an allen potenziellen Teilnehmenden des Online-Surveys war deutlich kleiner.

Die Zahl der Expert*innen, die über den Online-Survey in die Entwicklung der AGIL-Kriterien einbezogen wurden, ist als sehr positiv zu bewerten, der moderate Rücklauf von 31,0 % (195/630) als zufriedenstellend, zumal die Leitlinienexpert*innen den Leitlinienprozess i. d. R. ehrenamtlich unterstützen und die verfügbare Zeit ein großer limitierender Faktor ist.

Für die Repräsentativität der Stichprobe ist zu berücksichtigen, dass diese zwar bei einem aus unserer Sicht zufriedenstellenden Anteil von 31,0% der kontaktierten Leitlinienexpert*innen lag, im Umkehrschluss aber 69,0% nicht teilnahmen. 68,2% der Teilnehmenden an dem Survey waren Ärzt*innen, und 88,7% der Teilnehmenden haben bereits mindestens einmal an einem Leitlinien-Aktualisierungsprozess mitgewirkt. Der Umfang der von den Teilnehmenden erhobenen Charakteristika wurde bewusst begrenzt, um die Akzeptanz für die Teilnahme am Online-Survey zu erhöhen.

Die offenen Fragen wurden selten ausgefüllt (Range: n = 11 bis n = 33). Daher liegen für diese deutlich weniger Daten der Leitlinienexpert*innen vor als für die quantitativen Fragen des Online-Surveys (N = 195), die für die Definition „vollständig ausgefüllter Fragebogen“ herangezogen wurden. Ob eine Sättigung bezüglich der qualitativen Ergebnisse eingetreten ist, können wir daher nicht abschließend bewerten.

Das Versenden des Online-Surveys über ÄZQ, AWMF-IMWi und RKI erforderte die Erstellung eines einheitlichen Einladungslinks in SoSci-Survey, was eine mehrfache Teilnahme an dem Survey theoretisch möglich machte. Da es sich bei dem Online-Survey jedoch nicht um eine Wahl (z. B. von Personen oder Abstimmung eines Konsenses), sondern um eine Bewertung einer Kriterienliste handelte, das Ausfüllen etwas Zeit in Anspruch nahm und nur vollständig ausgewertete Fragebögen bezüglich der quantitativen Fragen gezählt wurden, gehen wir davon aus, dass ein Bias der Ergebnisse durch eine mehrfache Teilnahme eher unwahrscheinlich ist.

Es ist anzumerken, dass wir aus datenschutzrechtlichen Gründen keine persönlichen Daten der Workshopteilnehmenden erheben konnten. Wir gehen jedoch davon aus, dass in den Online-Survey vorwiegend das Wissen klinisch tätiger Leitlinienexpert*innen eingeflossen ist, während an dem Workshop im Rahmen der Jahrestagung des EbM-Netzwerks überwiegend wissenschaftlich tätige Personen mit Erfahrung in der Erstellung von Leitlinien teilnahmen, sodass sich beide Gruppen gut ergänzten.

Schlussfolgerung

Die AGIL-Kriterien wurden von erfahrenen Methodiker*innen und Wissenschaftler*innen aus dem Bereich der Leitlinienarbeit unter Einsatz von quantitativen und qualitativen Methoden in einem mehrstufigen Prozess entwickelt. Die Einbeziehung einer umfangreichen, erfahrenen Expert*innengruppe über den Online-Survey und den Workshop beim EbM-Kongress waren methodische Schritte dieses Projekts, die die Vollständigkeit, Wichtigkeit und Verständlichkeit der AGIL-Kriterien stärken.

Im Vergleich zu anderen Ansätzen besteht der Mehrwert der AGIL-Kriterien darin, dass sie kompakt und schnell anwendbar sind und sich der Blick bei Anwendung der AGIL-Kriterien in die Zukunft richtet. Dabei steht die prospektive Einschätzung des zukünftigen Aktualisierungsbedarfs einzelner Leitlinienempfehlungen zum Zeitpunkt der Fertigstellung einer Leitlinie oder Leitlinienaktualisierung im Fokus. Auf diese Weise leisten die AGIL-Kriterien einen Beitrag zu einer effizienten Priorisierung von evidenzbasierten Leitlinien- und Impfeempfehlungen und ermöglichen so eine vorausschauende Planung von Aktualisierungsprozessen.

In einem nächsten Schritt werden die aktuellen AGIL-Kriterien (Version 1.0) in ausgewählten Leitliniengruppen implementiert, evaluiert und ggf. angepasst.

Ethik

Für die hier berichtete Befragung wurde ein positives Votum der Ethikkommission des Universitätsklinikums Freiburg eingeholt (22-1150).

Verfügbarkeit der Daten

Die anonymisierten Daten des Online-Surveys können bei begründeten Anfragen zur Verfügung gestellt werden (waldemar.siemens@uniklinik-freiburg.de). Eine Word-Datei der AGIL-Kriterien 1.0 (siehe [Tabelle 2](#)) kann nach erfolgter Anfrage versendet werden.

Förderung

Das DEAL Projekt wurde vom Innovationsfonds des Gemeinsamen Bundesausschuss gefördert (Förderkennzeichen: 01VSF20019).

Danksagung

Wir danken Dr. Sebastian Voigt-Radloff (Institut für Medizinische Biometrie und Statistik, Sektion Versorgungsforschung und Rehabilitationsforschung, Universitätsklinikum Freiburg, Deutschland) für die Konzeption des Projektes, die Förderakquise und die Unterstützung in der frühen Phase der Planung des DEAL-Projektes. Wir danken Gina Bantle (Institut für Evidenz in der Medizin, Universitätsklinikum Freiburg, Deutschland) für die Mitgestaltung der Visualisierung der Ergebnisse. Frau Simone Witzel (AWMF-IMWi) danken wir für die Unterstützung beim Workshop des EbM Kongresses. Schließlich danken wir allen Leitlinienexpert*innen, die beim Online-Survey und beim EbM Workshop teilgenommen haben.

Interessenkonflikt

Die Autor*innen geben an, dass keine Interessenkonflikte vorliegen.

Autor*innenschaft

Waldemar Siemens: Konzeption des Manuskriptes, Methodik – Entwicklung des Online-Surveys, Methodik – Durchführung des EbM Workshops, Formale Analyse, Datenerhebung, Verfassen des Manuskripts – Erster Entwurf, Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Sonja Mahler: Methodik – Entwicklung des Online-Surveys, Formale Analyse, Datenerhebung, Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Corinna Schaefer: Konzeption des DEAL-Projekts, Methodik – Durchführung des EbM Workshops, Verfassen des Manuskripts – Begutachtung und Überarbeitung, Supervision.

Monika Nothacker: Konzeption des DEAL-Projekts, Methodik – Durchführung des EbM Workshops (Vorbereitung, in Präsenz vertreten durch Simone Witzel, AWMF-IMWi), Verfassen des Manuskripts – Begutachtung und Überarbeitung, Supervision, Förderakquise.

Vanessa Piechotta: Verfassen des Manuskripts – Begutachtung und Überarbeitung, Supervision.

Peggy Prien: Methodik – Durchführung des EbM Workshops, Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Sabine Schüler: Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Sabine Schwarz: Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Susanne Blödt: Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Iris Thielemann: Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Thomas Harder: Verfassen des Manuskripts – Begutachtung und Überarbeitung, Supervision, Förderakquise.

Philipp Kapp: Methodik – Entwicklung des Online-Surveys, Verfassen des Manuskripts – Begutachtung und Überarbeitung.

Valérie Labonté: Methodik – Entwicklung des Online-Surveys, Verfassen des Manuskripts – Begutachtung und Überarbeitung, Supervision.

Joerg J. Meerpohl: Konzeption des DEAL-Projekts, Methodik – Entwicklung des Online-Surveys, Verfassen des Manuskripts – Begutachtung und Überarbeitung, Supervision, Förderakquise.

Cordula Braun: Konzeption des DEAL-Projekts, Konzeption des Manuskripts, Methodik – Entwicklung des Online-Surveys, Methodik – Durchführung des EbM Workshops, Verfassen des Manuskripts – Begutachtung und Überarbeitung, Supervision, Förderakquise.

Anhang A. Zusätzliche Daten

Zusätzliche Daten verbunden mit diesem Artikel finden sich in der Online-Version unter: <https://doi.org/10.1016/j.zefq.2023.11.006>.

Literatur

- [1] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). AWMF-Regelwerk Leitlinien, <https://www.awmf.org/regelwerk/>, 2023 (aufgerufen am 15.06.2023).
- [2] Akl EA, Meerpohl JJ, Elliott J, Kahale LA, Schunemann HJ. Living systematic review N. Living systematic reviews: 4. Living guideline recommendations. J Clin Epidemiol 2017;91:47–53. <https://doi.org/10.1016/j.jclinepi.2017.08.009>.
- [3] Becker M, Neugebauer EA, Eikermann M. Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure. J Clin Epidemiol 2014;67:33–45. <https://doi.org/10.1016/j.jclinepi.2013.06.021>.
- [4] Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, et al. Living systematic review: 1. Introduction-the why, what, when, and how. J Clin Epidemiol 2017;91:23–30. <https://doi.org/10.1016/j.jclinepi.2017.08.010>.
- [5] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). S3-Leitlinie Empfehlungen zur Therapie von Patienten mit COVID-19 – Living Guideline, <https://register.awmf.org/de/leitlinien/detail/113-001LG>, 2022 (aufgerufen am 15.06.2023).
- [6] World Health Organization (WHO). Clinical management of COVID-19: Living guideline, <https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2023.1>, 2023 (aufgerufen am 15.06.2023).
- [7] Lamontagne F, Agarwal A, Rochwerf B, Siemieniuk RA, Agoritsas T, Askie L, et al. A living WHO guideline on drugs for covid-19. BMJ 2020;370. <https://doi.org/10.1136/bmj.m3379>.
- [8] Sanabria AJ, Pardo-Hernandez H, Ballesteros M, Canelo-Aybar C, McFarlane E, Nino de Guzman E, et al. The UpPriority tool was developed to guide the prioritization of clinical guideline questions for updating. J Clin Epidemiol 2020;126:80–92. <https://doi.org/10.1016/j.jclinepi.2020.06.018>.

- [9] Cheyne S, Fraile Navarro D, Buttery AK, Chakraborty S, Crane O, Hill K, et al. Methods for living guidelines: early guidance based on practical experience. Paper 3: selecting and prioritizing questions for living guidelines. *J Clin Epidemiol* 2023;155:73–83. <https://doi.org/10.1016/j.jclinepi.2022.12.021>.
- [10] Gemeinsamer Bundesausschuss (G-BA). DEAL – Dynamische Evidenzaktualisierung für Aktuelle Leitlinienempfehlungen, <https://innovationsfonds.g-ba.de/projekte/versorgungsforschung/deal-dynamische-evidenzaktualisierung-fuer-aktuelle-leitlinienempfehlungen.399>, 2021 (aufgerufen am 15.06.2023).
- [11] Braun CNM, Schaefer C, Labonté V, Siemens W, Mahler S, Meerpohl JJ. Entwicklung einer Kriterienliste für die prospektive Einschätzung des Aktualisierungsbedarfs von Leitlinienempfehlungen. *Gesundheit und Klima – EbM für die Zukunft 24 Jahrestagung des Netzwerks Evidenzbasierte Medizin*. Potsdam, Düsseldorf: German Medical Science GMS Publishing House; 2023.
- [12] Sharma A, Minh Duc NT, Luu Lam Thang T, Nam NH, Ng SJ, Abbas KS, et al. A consensus-based checklist for reporting of survey studies (CROSS). *J Gen Intern Med* 2021;36:3179–87. <https://doi.org/10.1007/s11606-021-06737-1>.
- [13] Kuckartz U. Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. Beltz Juventa; 2014.
- [14] Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:. <https://doi.org/10.1136/bmj.l4898>.
- [15] Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:. <https://doi.org/10.1136/bmj.j4008>.
- [16] Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is “quality of evidence” and why is it important to clinicians? *BMJ* 2008;336:995–8. <https://doi.org/10.1136/bmj.39490.551019.BE>.
- [17] Fraile Navarro D, Cheyne S, Hill K, McFarlane E, Morgan RL, Murad MH, et al. Methods for living guidelines: early guidance based on practical experience. Paper 5: decisions on methods for evidence synthesis and recommendation development for living guidelines. *J Clin Epidemiol* 2023;155:118–28. <https://doi.org/10.1016/j.jclinepi.2022.12.022>.
- [18] Cheyne S, Fraile Navarro D, Hill K, McDonald S, Tunnicliffe D, White H, et al. Methods for living guidelines: early guidance based on practical experience. Paper 1: Introduction. *J Clin Epidemiol* 2023;155:84–96. <https://doi.org/10.1016/j.jclinepi.2022.12.024>.