

Diplomarbeit

**Darstellung vorhandener Instrumente zur Bewertung
der Qualität von Gesundheitsinformationen**

eingereicht von

Isabelle Yolanda Beck-Werz

zur Erlangung des akademischen Grades

**Doktorin der gesamten Heilkunde
(Dr. med. univ.)**

an der

Medizinischen Universität Graz

ausgeführt am

**Institut für Allgemeinmedizin und evidenzbasierte
Versorgungsforschung (IAMEV)**

unter der Anleitung von

Univ. Prof. Dr. med. univ. Andrea Siebenhofer-Kroitzsch

Dr. med. univ. Nicole Posch

Graz, am 12.11.2021

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am 12.11.2021

Isabelle Yolanda Beck-Werz eh

Inhaltsverzeichnis

Abkürzungen und deren Erklärung.....	3
Abbildungsverzeichnis	4
Tabellenverzeichnis	5
Zusammenfassung in Deutsch	6
Abstract in Englisch	7
Einleitung	8
1.1 Die Informationsflut	8
1.2 Vor- und Nachteile der wachsenden Anzahl von Gesundheitsinformationen.....	8
1.3 Qualitätskriterien für gute Gesundheitsinformationen	9
1.4 Gesundheitskompetenz in Österreich	11
1.5 Unterschied zwischen Gesundheitsinformation und Entscheidungshilfe.....	13
1.6 Instrumente zur Bewertung von GI	13
2 Material und Methoden	14
3 Ergebnisse – Resultate mit graphischen Darstellungen.....	15
3.1 Ergebnisse Anwenderstudien	18
3.2 Ergebnisse Reviews	19
3.3 Ergebnisse Originalpublikationen	20
3.4 Beschreibung der einzelnen Instrumente.....	22
3.4.1 DISCERN	22
3.4.2 EQIP 36	23
3.4.3 HONcode	25
3.4.4 IPDASi v4.0	26
3.4.5 PEMAT.....	28
3.4.6 LIDA.....	29
3.4.7 Suitability of Materials (SAM).....	31
3.5 Beschreibung der Instrumente zur Beurteilung der Lesbarkeit.....	31
3.5.1 Flesch Reading Ease Score (FRES).....	32
3.5.2 Flesch Kinkaid Reading Grade (FKGL).....	32
3.5.3 Gunning Fog Index (GFI).....	33
3.5.4 Simple Measure of Gobbledygook for readability (SMOG).....	33

3.6	Gegenüberstellung der Instrumente/ Kriterien	33
3.6.1	In sechs Instrumenten verwendetes Kriterium	34
3.6.2	In fünf Instrumenten verwendetes Kriterium	34
3.6.3	In vier Instrumenten verwendete Kriterien.....	34
3.6.4	In drei Instrumenten verwendete Kriterien.....	35
3.6.5	In zwei Instrumenten verwendete Kriterien	36
3.7	Überprüfung der Instrumente auf die 15 Qualitätskriterien	41
4	Diskussion	43
	Literaturverzeichnis	47

Abkürzungen und deren Erklärung

GI	Gesundheitsinformationen
IAMEV	Institut für Allgemeinmedizin und evidenzbasierte Versorgungsforschung
PDA	patient decision aids
ÖPGK	Österreichische Plattform für Gesundheitskompetenz

Abbildungsverzeichnis

Abbildung 1 Gesundheitskompetenz nach Land und für die Gesamtstichprobe	12
Abbildung 2 Flowchart Literaturrecherche	17

Tabellenverzeichnis

Tabelle 1 Auswertung Anwenderstudien.....	18
Tabelle 2 Anzahl Instrumente in den Reviews	20
Tabelle 3 Häufigkeit Instrumente in den Reviews.....	20
Tabelle 4 Ergebnisse Originalpublikationen.....	22
Tabelle 5 Einteilung des FRES	32
Tabelle 6 Auflistung Qualitätskriterien	37
Tabelle 7 Überprüfung der 15 Qualitätskriterien in den Instrumenten zur Bewertung der Qualität der GI.....	42

Zusammenfassung in Deutsch

Hintergrund Die Anzahl von Gesundheitsinformationen als Print-Material, sowie im Internet steigt stetig. Dies erschwert es den Patientinnen/ Patienten die Qualität dieser zu beurteilen. Deswegen wurden zur Unterstützung viele Instrumente zur Bewertung der Qualität von verschiedenster Gesundheitsinformation entwickelt. Da immer mehr Instrumente entwickelt werden, verschiebt sich das Grundproblem um eine Stufe, da sich nun für Patientinnen/ Patienten die Frage stellt, welches Instrument valide ist. Um einen Überblick dieser Zahl von Instrumenten zu behalten, wurden in dieser Arbeit die Instrumente übersichtlich dargestellt.

Methodik Zuerst wurde eine systematische Literaturrecherche durchgeführt und mit den Ergebnisse dieser ein Screening erstmal auf Abstract- und anschließend auf Volltextebene durchgeführt. Daraufhin wurden die Anwenderstudien und die Reviews auf verwendete oder genannte Instrumente zur Beurteilung der Qualität von Gesundheitsinformationen durchsucht. Im Anschluss wurden die gefundenen Instrumente und die Instrumente aus den Originalpublikationen erneut bezüglich der Ausschlusskriterien untersucht. Von den Instrumenten DISCERN, EQIP36, IPDASi, HON, LIDA, SAM und PEMAT wurden die Kriterien genauer betrachtet und miteinander verglichen.

Ergebnisse Bei der systematischen Literaturrecherche wurden insgesamt 2425 verschiedene Publikationen gefunden. Nach dem Screening auf Abstractebene verblieben noch 190 Publikationen und nach dem Volltextscreening 174. Von den 174 Publikationen handelte es sich bei 8 Stück um Originalpublikationen, bei 13 um Reviews und bei 153 um Anwenderstudien. Bei genauerer Betrachtung der Anwenderstudien wurden insgesamt 62 unterschiedliche Instrumente gefunden. In den 13 Reviews wurden 121 verschiedene Instrumente zur Sprache gebracht. Die 8 Originalpublikationen behandeln verschiedene Versionen von den Instrumenten EQIP, IPDASi, DISCERN und PEMAT.

Schlussfolgerung In dieser Arbeit wurden eine Vielzahl von Instrumenten gefunden. Die meisten sind aber aufgrund von unterschiedlichsten Gründen im praktischen Alltag ungeeignet.

Abstract in English

Background The amount of health information available in print and on the Internet is constantly increasing. This makes it more difficult for patients to assess the quality of this information. Therefore, many instruments have been developed to help assess the quality of various health information. As more and more instruments are developed, the basic problem is shifted one step further, as patients are now faced with the question of which instrument is valid. To maintain an overview over the number of Instruments, the instruments were clearly presented in this paper.

Methods At first there has been a systematic literature search, followed by a screening of the results, carried out on an abstract level and later on the full text. Thereupon, the user studies and the reviews were searched for instruments used or mentioned to assess the quality of health information. After that the instruments, which have been found, and the instruments from the original publication have been examined again through the exclusion criteria. The criteria of the instruments DISCERN, EQIP36, IPDASi, HON, LIDA, SAM, and PEMAT have been studied precisely and have been compared to each other.

Results A total of 2425 different publications were found through the systematic literature search. After the screening at abstract level, only 190 publications remained, and after full text screening there were 174 left. 8 of the 174 publications were original publications, 13 were reviews, and 153 were user studies. A closer look at the user studies revealed a total of 62 different instruments. In the 13 reviews, 121 different instruments were brought up. The 8 original publications deal with different versions of the instrument EQIP, IPDASi, DISCERN, and PEMAT.

Conclusion A large number of instruments were found in this work. However, most of them are unsuitable for practical use due to a wide variety of reasons.

Einleitung

1.1 Die Informationsflut

Die Anzahl an medizinischen Informationen steigt stetig - vor allem digitale Angebote, aber auch analoge. Beispielsweise werden bei einer Suchanfrage in Google mit dem Suchbegriff „Diabetes mellitus“ am 11.3.2020 um 16:18 Uhr 37 700 000 Treffer angezeigt und ungefähr eineinhalb Jahre später am 02.09.2021 um 19:57 Uhr 38 700 000 Treffer. Somit ist die Anzahl der Treffer in den eineinhalb Jahren um eine Millionen gestiegen. Daraus folgt, dass die Anzahl von Websites, die über medizinische Themen informieren, immer mehr zu nehmen. Die Bevölkerung ist somit mit einer wachsenden Informationsflut konfrontiert.

Aber nicht nur die Anzahl der Gesundheitsinformationen (GI) steigt stetig, sondern auch die Nachfrage danach. Laut einer Umfrage der Bertelsmann Stiftung im Jahr 2017 (1) haben nur 12 Prozent der 1074 Befragten angegeben sich in den letzten 12 Monate nicht über Gesundheitsthemen informiert zu haben. Ganze 88 Prozent haben sich andererseits über Gesundheitsthemen informiert. Diese Zahl habe sich laut Umfragen des Gesundheitsmonitors seit 2001 mehr als um ein Drittel erhöht (1). Dies zeigt den steigenden Trend der Nachfrage nach GI. Dabei werden klassische Informationsquellen wie Zeitschriften immer noch am häufigsten verwendet, aber besonders bei spezifischen Gesundheitsfragen wird vor allem das Internet zu Rate gezogen (1). Wie 2018 eine weitere Umfrage der Bertelsmann Stiftung (2) herausgefunden hat, wird die Internetsuche dabei meist mit der Suchmaschine Google begonnen. Des Weiteren beschreibt die Umfrage, dass das Vertrauen in Google sehr groß ist und die Ergebnisse als „überwiegend gut verständlich“ (2) wahrgenommen werden.

1.2 Vor- und Nachteile der wachsenden Anzahl von Gesundheitsinformationen

Die Option jederzeit an eine Vielzahl von GI zu gelangen und sich selbständig informieren zu können hat einige Vorteile. Es wird angenommen, dass die Zahl an verspäteten Diagnosen durch die eigenständige Suche von Patientinnen/ Patienten vermindert werden kann (3). Außerdem wirkt sich die Möglichkeit, schon vor dem Arztbesuch immer leichter an Gesundheitsinformationen zu gelangen und das große Vertrauen in Google nicht zwangsläufig negativ auf Arzt-Patienten-Beziehung aus, sondern diese kann dadurch sogar verbessert werden – auch wenn man anfangs das Gegenteil vermuten mag. Dabei ist vor

allem wichtig, dass die Patientin/ der Patient die gefundenen Informationen mit der Ärztin/ dem Arzt bespricht (4). Auch gerade bei der Erstdiagnose einer chronischen Erkrankung kann die zusätzliche selbstständige Informationssuche ein wichtiger Bestandteil sein, um sich auf die Erkrankung leichter einstellen zu können und die Krankheit leichter in den Alltag integrieren zu können (2). Dabei kann vor allem in Selbsthilfenetzwerken oder -foren Rückhalt gefunden werden.

Es kommt jedoch zu Problemen, wenn der Informationsgehalt nicht vollständig, unabhängig und korrekt ist. Dabei kann es zu einer Fehlinformation kommen. Bei Medienberichten ist die Evidenz der Informationen zum Beispiel meistens stark verzerrt. Laut einer Studie von Kerschner et al. (5) geben nur 10,8% der 990 in die Studie einbezogenen Medienberichte die tatsächliche Studienlage richtig wieder und verbreiten somit korrekte und evidenzbasierte Informationen. Dabei macht es auch keinen Unterschied, ob es sich um ein Online oder Print Medium handelt.

Im MEDIB-Projekt (medizinische Informationsbroschüren in der Hausarztpraxis) (6) wurde herausgefunden, dass die Informationsmaterialien in Hausarztpraxen in der Steiermark zum Großteil nicht den Qualitätsstandards für gute Gesundheitsinformationen entsprechen. In diesem Projekt wurden Informationsmaterialien mit Hilfe des EQIP- Instruments (Ensuring Quality Information for Patients) (7) bewertet. Dabei erreichten die Informationsmaterialien durchschnittlich nur eine Bewertung von 39 von insgesamt 100 möglichen Prozentpunkten.

1.3 Qualitätskriterien für gute Gesundheitsinformationen

Um diesen Schwierigkeiten entgegenzuwirken hat die Österreichische Plattform Gesundheitskompetenz (ÖPGK) in Kooperation mit dem Bundesministerium für Gesundheit und Frauen (BMGF) auf Basis der guten Praxis Gesundheitsinformation 2.0 des Deutschen Netzwerkes für evidenzbasierte Medizin e.V. (8) eine Richtlinie für das Erstellen von guten Gesundheitsinformationen entwickelt, in der sie 15 Qualitätskriterien für gute GI beschrieben haben.

Diese 15 Kriterien sind in drei Kategorien - Grundlagen, Auswahl und Darstellung der Fakten und Glaubwürdigkeit – unterteilt. Zu jedem Kriterium wurde ein Ziel definiert und wichtige Fragen erstellt, welche bei der Auswertung einer GI unterstützen. Zu der Kategorie „Grundlagen“ gehören drei Kriterien. Das erste Kriterium ist die „Identifizierung besonderer Informationsbedürfnisse“ (8). Dabei soll die genaue Zielgruppe definiert werden, um daraufhin prüfen zu können, ob die GI für diese passend ist. Bei dem zweiten Kriterium geht

es um die „systematische Recherche“ (8), bei der die Quellen vollständig und ausgeglichen bezüglich der Relevanz für die Zielgruppen und überprüfbar sein sollen. Die Aktualität bezüglich der Quellen bzw. Evidenzen wird im dritten Qualitätskriterium besprochen. Dabei ist das Ziel, dass „die Zielgruppe [...] aktuellste Informationen aus den besten verfügbaren Studien“ (8) erhält. Verzerrungen und Wissenslücken sollen beschrieben sein.

Die Kategorie „Auswahl und Darstellung der Fakten“ enthält acht Qualitätskriterien. Dabei sollen laut dem ersten Kriterium in dieser Kategorie die Endpunkte beschrieben werden, damit eingeschätzt werden kann, wie die „Auswirkungen auf [das] alltägliche Leben“ (8) aussehen bzw. ob die berichteten Ergebnisse relevant für die jeweilige Zielgruppe sind. Das nächste Kriterium ist die „Wahl und Darstellung von Vergleichen“ (8). Dazu gehört, dass einerseits die verschiedenen Optionen und andererseits Nutzen und Risiken jeder Option genannt und beschrieben werden sollten, wie auch das Abwarten oder Nichtstun, damit eine individuelle Entscheidung unterstützt werden kann. Beim Kriterium Umgang mit Zahlen und Risikoangaben ist eine „ausgewogen[e], geschlechtergerecht[e] und verständlich dargestellt[e]“ (8) Darstellung von Nutzen und Risiko das Ziel. Im nächsten Kriterium geht es um die „Berücksichtigung von Alters- und Geschlechterunterschieden“ (8) bei den Informationen bezüglich Nutzen und Risiko, natürlichem Krankheitsverlauf und Wirkung und unerwünschte Wirkung von Interventionen. Die wichtige Frage hier ist, „wie [sich] diese Informationen nach Alter, Geschlecht, Lebenssituation und Begleitumständen“ (8) unterscheiden. Das achte Kriterium hat das Ziel „Inhalt, kulturelle Aspekte, Sprache, Gestaltung und Medium [...] auf die Zielgruppe“ (8) abzustimmen. Die Darstellung der GI soll sachlich angemessen sein und nicht dramatisieren oder verharmlosen. Außerdem soll die GI eindeutig zwischen Information und Empfehlung unterscheiden. „Werden außer Fakten auch Empfehlungen gegeben?“ (8) und „Wie werden die Empfehlungen klar erkennbar gemacht?“ (8) sind bei diesem Kriterium wichtige Fragen. Das elfte Qualitätskriterium heißt „Vorgehen bei der Erstellung von Entscheidungshilfen“ (8). Dabei sollte das Ziel sein, dass mit Hilfe der GI eine Entscheidung bezüglich der Untersuchungs- und Behandlungsmöglichkeit gefällt werden kann. Die letzte Kategorie „Glaubwürdigkeit“ besteht aus vier Qualitätskriterien. Im ersten Kriterium hier geht es darum, ob die Namen der Personen, welche die GI erstellt und finanziert haben, klar ersichtlich sind, und ob der Datenschutz vor allem bei interaktiven Webseiten gesichert ist. Beim 13. Qualitätskriterium geht es um die Interessenskonflikte, welche hinter der Erstellung jeglicher GI stecken können. Bei diesem Kriterium wird die Frage gestellt, „welche finanziellen oder

persönlichen Interessen oder Abhängigkeiten [...] Personen oder Organisationen [haben], die die Informationen herausgeben/ erstellen“ (8). Um den potenziellen Nutzerinnen/ Nutzer eine schnelle Entscheidung zu ermöglichen, ob die GI die gewünschten Informationen enthält, soll eine Kurzbeschreibung der Inhalte vorhanden sein. Das letzte Kriterium dieser fünfzehn Qualitätskriterien beschäftigt sich mit der Aktualität von GI. Dabei soll „die Aktualität und de[r] Aktualisierungsprozess der Informationen“ (8) anhand eines Erscheinungsdatums und Überarbeitungsdaten beurteilt werden können.

Diese Qualitätskriterien könnten möglicherweise dazu beitragen, dass die Bevölkerung über Gesundheitsthemen gut informiert wird. Aber auch wenn bereits viele GI mit Hilfe dieser Richtlinie erstellt werden, sind noch viele GI erhältlich, die nicht den Qualitätsansprüchen entsprechen. Deswegen ist es zusätzlich wichtig, dass die Bevölkerung gute GI aus der Informationsflut filtern kann.

1.4 Gesundheitskompetenz in Österreich

Um differenzieren zu können, ob es sich um eine gute oder schlechte GI handelt, ist die Gesundheitskompetenz einer Person selbst entscheidend. Der Begriff „Gesundheitskompetenz“ oder im Englischen „health literacy“ ist ein komplexer und vielschichtiger Begriff. Von der EU zum Beispiel wird „Health literacy“ als Fähigkeit beschrieben, GI zu lesen, zu filtern und zu verstehen, um fundiert urteilen zu können (7). Zusätzlich sind noch viele verschiedene weitere Definitionen im Umlauf.

Für eine einheitliche Definition wurden 2017 im Rahmen eines systematischen Reviews 17 verschiedene Definitionen zusammengetragen und eine neue Definition formuliert “Health literacy is linked to literacy and entails people’s knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course” (9). Diese Definition beinhaltet die vier Stufen der Kompetenz, welche im Review definiert wurden.

Diese vier Stufen beschreiben die Fähigkeit...

1. ...GI zu suchen, zu finden und zu erhalten.
2. ...GI zu verstehen.
3. ...GI zu interpretieren, zu filtern, zu beurteilen und zu evaluieren.

4. ...mit der GI Entscheidungen zur Erhaltung oder Verbesserung der Gesundheit fällen zu können (9).

Die Gesundheitskompetenz in Europa ist laut dem European Health Literacy Survey (HLS-EU) sehr niedrig (10). Dies wurde 2015 im Rahmen einer Umfrage festgestellt. Mit dem HLS-EU-Q-Instrument wurde die Gesundheitskompetenz von acht europäischen Ländern gemessen. Man fand heraus, dass jede zweite befragte Person der acht Länder, eine begrenzte Gesundheitskompetenz besitzt. Aber vor allem Österreich schnitt im europäischen Vergleich sehr schlecht ab (10). Der Anteil der befragten Personen, die eine inadäquate oder problematische Gesundheitskompetenz haben, lag bei 56% in Österreich (Abbildung 1). Nur noch Bulgarien mit einem Anteil von 62% und Spanien mit 58% schnitten noch schlechter ab. Im Vergleich der Mittelwerte des Gesundheitskompetenz-Index lag Österreich mit einem Index von 32 sogar auf dem vorletzten Platz (9).

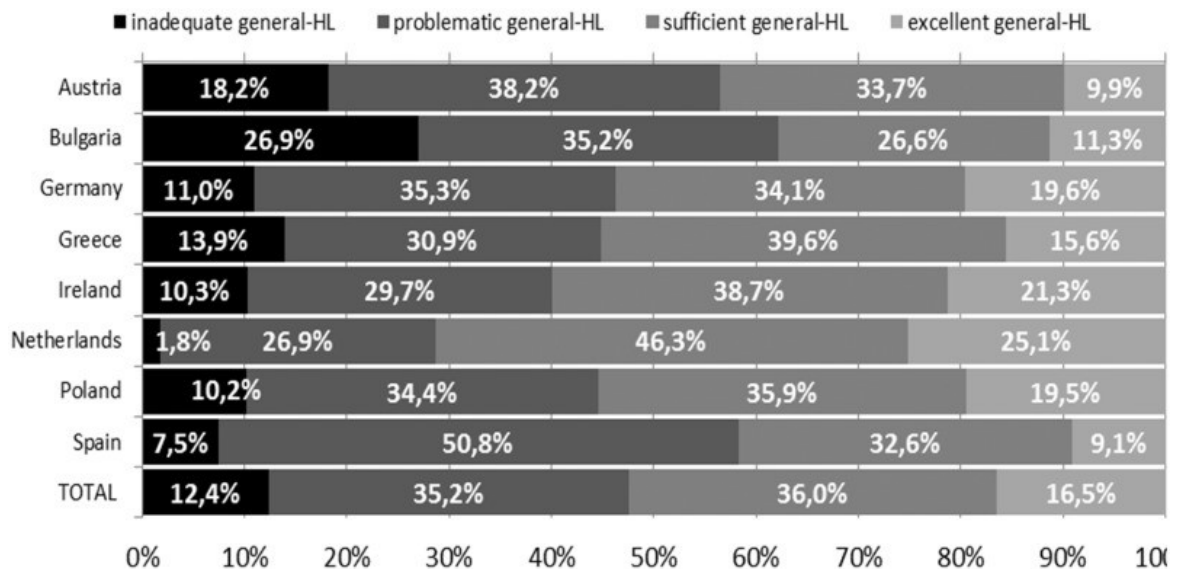


Abbildung 1 Gesundheitskompetenz nach Land und für die Gesamtstichprobe

Der Zustand der schlechten Gesundheitskompetenz von Österreich ist der Regierung bewusst. Im aktuellen Zielsetzungsvertrag des Bundesministerium für Gesundheit und Frauen von 2017-2021 (11) ist die Steigerung der österreichischen Gesundheitskompetenz als Ziel fest verankert. Dieser Zielsetzungsvertrag ist nun auch stärker mit den Gesundheitszielen Österreichs verbunden, welche ebenfalls vom Bundesministerium für Gesundheit und Frauen entwickelt wurden. Eines der zehn Gesundheitsziele ist nun die Stärkung der Gesundheitskompetenz der Bevölkerung (12). Für die Erfüllung des Ziels wurden drei Maßnahmen als besonders wichtig eingestuft.

Diese sind...

- ...das Stärken der persönlichen Kompetenzen bei allen Bevölkerungsgruppen.
- ...den Zugang zu verständlichen Informationen zu erleichtern.
- ...das Fördern der Rolle der Patientinnen/ Patienten als gleichberechtigte Partnerinnen/ Partner (12).

1.5 Unterschied zwischen Gesundheitsinformation und Entscheidungshilfe

Es gibt verschiedenste Arten von Informationen, deren Grenzen oft verschwimmen. Dabei ist die Differenzierung von Gesundheitsinformationen und Entscheidungshilfen relevant. Eine Arbeitsgruppe des Deutschen Netzwerks evidenzbasierter Medizin hat eine Hilfestellung zur Erstellung von guten Gesundheitsinformationen entwickelt und dabei auch die Begrifflichkeiten definiert. Gesundheitsinformationen betreffen, laut des Deutschen Netzwerks evidenzbasierter Medizin, Gesundheitsthemen, wie zum Beispiel das grundlegende Wissen über Krankheit und Gesundheit, Maßnahmen zu den vier Präventionsformen, Diagnostik, Behandlung, Palliation und Nachsorge. Ebenfalls werden Themen zur Pflege und Krankheitsbewältigung abgedeckt (13).

Nicht zu verwechseln mit GI sind Entscheidungshilfen: Das sind Hilfsmittel, die der Patientin/ dem Patienten beim Abwägen verschiedener Optionen helfen sollen und der Patientin/ dem Patienten somit die Möglichkeit geben bei individuellen Entscheidungen, wie zum Beispiel bei der Therapie oder Diagnostik, mitbestimmen zu können (13). Dabei sollen Entscheidungshilfen nicht zu einer bestimmten Option raten oder eine Möglichkeit vorziehen und sollen auch kein Arztgespräch ersetzen, sondern der Patientin/ dem Patienten die Möglichkeit einer Partizipation bei der Entscheidung bieten (14). Sowohl Gesundheitsinformationen als auch Entscheidungshilfen sind in digitaler, als auch in analoger Form in unterschiedlichsten Medien vorhanden.

1.6 Instrumente zur Bewertung von GI

Mit Hilfe von Instrumenten zur Beurteilung der Qualität von GI kann die Differenzierung von guter und schlechter GI für die Bevölkerung erleichtert werden. Solche Instrumente bewerten GI nach bestimmten Kriterien und schätzen damit ihre Qualität ein. Das Angebot an Instrumenten für die Bewertung von GI steigt dabei ständig. Durch die zunehmende Zahl von Instrumenten, verschiebt sich damit das grundlegende Problem um eine Ebene: Wie weiß die Bevölkerung, welches Instrument überhaupt verlässlich ist? Viele Instrumente

werden unter nicht wissenschaftlichen Umständen, wie zum Beispiel von Einzelpersonen, entwickelt. Zudem wurden die Kriterien der einzelnen Instrumente noch nie miteinander verglichen und aufeinander abgestimmt.

Derzeit wird eine Checkliste mit 19 Kriterien (MAPPinfo) (15) erstellt und validiert, die auf den evidenzbasierten Kriterien der Leitlinien für gute Gesundheitsinformation basieren (8).

2 Material und Methoden

Zur Identifikation von bereits bestehenden Instrumenten zur Bewertung von GI wurde eine systematische Literaturrecherche durchgeführt. Über diese wurden Publikationen ausfindig gemacht, die Instrumente zur Bewertung von Gesundheitsinformationen enthalten. Diese Suche wurde am 10.07.2019 in der Datenbank Medline und am 18.07.2019 in der Datenbank PsycINFO durchgeführt. Grundlegend wurde nach Publikationen gesucht, welche die Themen „Patienteninformation“, „Entscheidungshilfe“ oder „Gesundheitsinformation“ in Kombination mit den Themenblock „Qualitätsbewertung“ und „Kriterien“, beinhalten. Dabei wurden unterschiedliche Kombinationen von MeSHterms und Textworten verwendet (siehe Suchstrategie im Anhang). Um einen möglichst vollständigen Überblick über bereits vorhandene Instrumente zur Bewertung von GI zu erhalten, wurden alle Publikationen bis zum Durchführungszeitpunkt einbezogen und keine weiteren Filter angelegt.

Die hierbei identifizierten Abstracts wurden von zwei Reviewern unabhängig voneinander nach deren Inhalten auf Abstractebene kategorisiert. Jeder Abstract wurde einer von drei Gruppen zugeordnet: Gruppe der Originalpublikationen, der Reviews oder der Anwenderstudien. Zusätzlich wurden die Instrumente zur Bewertung von GI, welche schon im Abstract mit Namen aufgetaucht sind, mit einem Kürzel versehen und in eine eigene Excel-Tabelle aufgenommen. Die Anwenderstudien wurden zusätzlich mit einem „Z“ versehen, falls hier eigene Kriterien oder Instrumente zur Qualitätsbewertung entwickelt wurden. Außerdem wurden jene Anwenderstudien mit einem Fragezeichen gekennzeichnet, bei denen nicht alle verwendeten Instrumente schon im Abstract benannt wurden. Wenn im Abstract ersichtlich wurde, dass das Paper inhaltlich nichts mit der Bewertung von Gesundheitsinformationen oder Entscheidungshilfen zu tun hat, wurde es sofort ausgeschlossen. Falls die beiden Reviewer zu einer unterschiedlichen Einschätzung des Abstracts hinsichtlich der Kategorisierung gekommen sind, hat ein dritter Reviewer das Abstract noch einmal unabhängig eingeordnet.

Dem Screening auf Abstractebene folgte ein Screening auf Volltextebene, wiederum durch zwei Reviewer unabhängig voneinander. Bei einer unterschiedlichen Einschätzung wurde diese wieder durch einen dritten Reviewer aufgelöst. Der Ausschluss folgte vorab klar definierten Ausschlusskriterien.

Insgesamt wurden acht Ausschlusskriterien vorab vom Projektteam definiert. Zum einen wurden Volltexte ausgeschlossen, wenn sie kein Manual oder Instrument für die Bewertung von GI oder Entscheidungshilfen enthielten. Des Weiteren mussten einzelne Kriterien zur Bewertung dieser klar erkennbar sein. Das dritte Ausschlusskriterium beinhaltete, dass es sich ausschließlich um eine Sammlung von Kriterien und nicht um ein Instrument handelte. Dies lag auch vor, wenn keine klare Methodik aus dem Volltext hervorging. Als nächstes wurden Paper ausgeschlossen, welche nicht auf Englisch oder Deutsch geschrieben waren. Das fünfte Ausschlusskriterium schloss Paper aus, wenn die Checkliste nur auf ein rein lokales Setting ausgerichtet war. Dieses Ausschlusskriterium traf zudem zu, wenn eine Einzelperson bei der Erstellung beteiligt war bzw. es nicht erkennbar war, dass es mehrere Personen waren. Das nächste Ausschlusskriterium war, dass Volltexte ausgeschlossen wurden, wenn diese Kriterien in Anwenderstudien entwickelt wurden. Wenn die Checkliste nicht frei verfügbar oder nicht mehr aufrufbar war, traf das siebte Ausschlusskriterium zu. Das letzte Ausschlusskriterium behandelte die Problematik, wenn Instrumente nur auf einen spezifisches Themenfeld, wie zum Beispiel eine spezielle Krankheit, ausgerichtet war. Dann wurden diese Volltexte ebenfalls ausgeschlossen.

Die Originalpublikationen, die im Abstractscreening als solche identifiziert wurden, wurden eingehend überprüft, ob sie den Ausschlusskriterien standhalten und dann mit einem jeweiligen Kürzel für das beschriebene Instrument versehen. Bei den Reviews wurden alle Instrumente zur Bewertung von GI, welche im Volltext genannt wurden, einzeln auf die Ausschlusskriterien überprüft und in einer weiteren Tabelle genau nach den jeweils beschriebenen Instrumenten aufgespalten. Auch die Volltexte der Anwenderstudien wurden im Folgenden nach beinhaltenden Instrumenten zur Bewertung von GI gesichtet, falls diese noch nicht im Abstract ersichtlich waren.

3 Ergebnisse – Resultate mit graphischen Darstellungen

Insgesamt wurden 2425 Publikationen im Rahmen der systematischen Literaturrecherche gefunden. Davon wurden 1876 in Medline auffindig gemacht und 549 in PsycINFO.

Im Rahmen des Abstractscreenings wurden insgesamt 2235 Publikationen exkludiert, wobei 1690 aus der Medline und 545 aus der PsycINFO Datenbank stammten. Auf 190 Publikationen traf kein Ausschlusskriterium zu und sie blieben dadurch inkludiert. Davon waren 26 Originalpublikationen, elf Reviews und 153 Anwenderstudien.

Bei der Vorbereitung des Volltextscreenings wurde eine weitere Publikation exkludiert, da für diese kein Volltext verfügbar war. Außerdem wurden zwei Publikationen, welche zuvor als potenzielle Originalpublikationen eingestuft wurden, in diesem Schritt nachträglich den Reviews zugeordnet.

Bei den Anwenderstudien wurden die Volltexte nach den beinhaltenden Instrumenten zur Bewertung von GI durchsucht, wenn diese nicht schon im Abstract genannt wurden. Zusätzlich wurden die Anwenderstudien nicht nach den Ausschlusskriterien gesichtet und somit wurde auch keine Anwenderstudie im Rahmen des Volltextscreening ausgeschlossen. Die verbleibenden 36 Volltexte wurden von zwei Reviewern unabhängig voneinander auf die acht vorab definierten Ausschlusskriterien überprüft. Bei den Reviews wurden alle im Volltext beschriebenen Instrumente zur Bewertung von Gesundheitsinformationen einzeln auf die Ausschlusskriterien überprüft. 15 Publikationen - potenzielle Originalpublikationen von Medline - wurden dabei auf Volltextebene exkludiert. Dabei traf am häufigsten das Ausschlusskriterium zu, dass es kein Manual oder Instrument für die Erstellung oder Bewertung von GI und/ oder eine Entscheidungshilfe beinhaltet.

Nach diesem Schritt wurden 174 Publikationen als relevant erachtet. Davon stammten 170 aus der Medline und vier aus der Psyc-INFO Datenbank. Insgesamt wurden 94 Instrumente zur Bewertung von GI detektiert.

Von den 174 Publikationen wurden acht als Originalarbeit, 13 als Reviews und 153 als Anwenderstudien identifiziert.

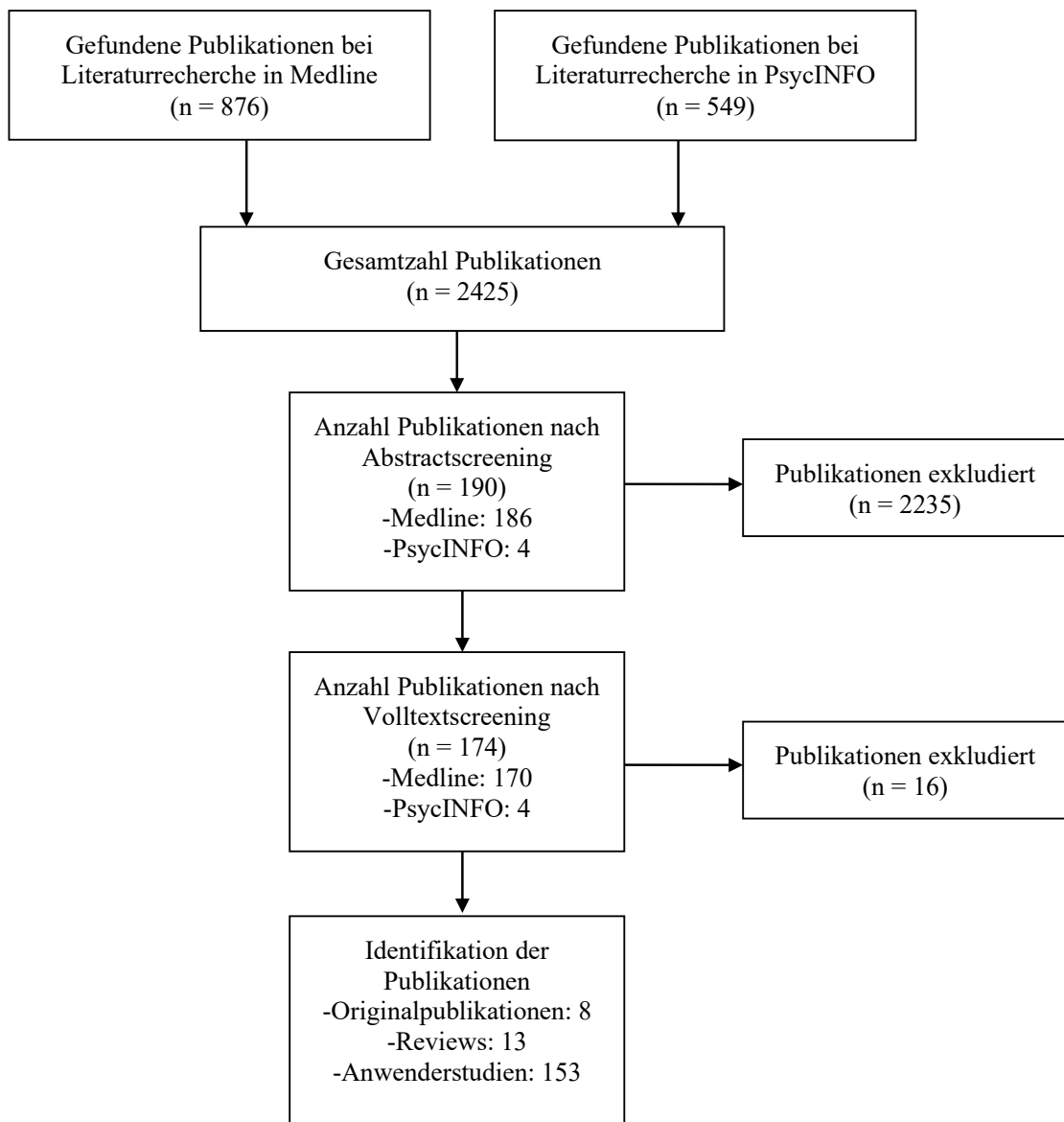


Abbildung 2 Flowchart Literaturrecherche

3.1 Ergebnisse Anwenderstudien

In der systematischen Literaturrecherche wurden 153 Publikationen als Anwenderstudien identifiziert. In 36 Fällen davon wurden eigene Kriterien oder Instrumente zur Qualitätsbewertung entwickelt. Diese selbstentwickelten Kriterien oder Instrumente wurden nicht genauer betrachtet, da bei diesen keine valide Erstellung der Kriterien oder Instrumente zu erwarten waren. In den 153 Anwenderstudien wurden insgesamt 62 unterschiedliche Instrumente zur Bewertung von Gesundheitsinformationen eingesetzt. Davon wurden 37 Instrumente nur in jeweils einer Publikation verwendet und acht jeweils in zwei Publikationen. Am häufigsten wurde das DISCERN Instrument (16) verwendet: 77 Mal und somit ungefähr in jeder zweiten Anwenderstudie wurde DISCERN zur Qualitätsbewertung benutzt. Auf dem zweiten Rang liegt die Quality Checklist (17), welche in 37 Anwenderstudien verwendet wurde. Des Weiteren wurde der Flesch Kinkaid Reading Grade (FKRG) (18) 29 Mal, HON (19) 28 Mal und der Flesch Reading Ease Score (FRES) (20) 25 Mal angewendet. Außerdem wurden die JAMA benchmarks/ standards (21) in 19, das Instrument LIDA (22) in 17 und SMOG (23) in 14 verschiedenen Anwenderstudien verwendet.

Die restlichen Instrumente werden nicht weiter aufgezählt, da diese in weniger als zehn Anwenderstudien behandelt wurden.

Instrument	Häufigkeit der Verwendung der Instrumente in den 153 Anwenderstudien
DISCERN	77
Quality Checklist	37
Selbst erstellte Kriterien	35
Flesch Kinkaid Reading Grade	29
HON	28
Flesch Reading Ease Score	25
JAMA benchmarks/ standards	19
LIDA	17
SMOG	14

Tabelle 1 Auswertung Anwenderstudien

3.2 Ergebnisse Reviews

In der systematischen Literaturrecherche wurden 13 Reviews identifiziert. Für jedes Review wurde eine Liste mit den beinhaltenden Instrumenten zur Bewertung von GI erstellt. Am meisten verschiedene Instrumente beinhaltete das Review von Breckons et al. (24) mit 37 verschiedenen Instrumente zur Bewertung von GI. 28 Instrumente zur Bewertung von GI bringen Luk et al. in ihrem Review (25) zur Sprache.

Autor und Titel des Reviews	Anzahl Instrumente
Breckons et al.: What do evaluation instruments tell us about the quality of complementary medicine information on the internet?	37
Luk et al.: Tools used to evaluate written medicine and health information: Document and user perspectives	28
Shedlosky-Shoemaker et al.: Tools for assessing readability and quality of health-related web sites	22
Hanif et al.: The role of quality tools in assessing reliability of the internet for health information	15
Daraz et al.: Can patients trust online health information? A meta-narrative systematic review addressing the quality of health information on the internet	14
Reavley et al.: The quality of mental disorder information websites: A review	13
Eysenbach et al.	11
Gagliardi et al.: Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination”	11
Cassidy et al.: Orthopaedic patient information on the world wide web: an essential review	10
Nouri et al.: Criteria for assessing the quality of mHealth apps: a systematic review	7
Huang et al.: Creating informed consumers and achieving shared decision making	7
Fahy et al.: Quality of patient health information on the internet: reviewing a complex and evolving landscape	5

Goto et al.: Oncology information on the internet	3
---	---

Tabelle 2 Anzahl Instrumente in den Reviews

Insgesamt werden in den 13 Reviews 121 verschiedene Instrumente zur Bewertung von GI genannt. Wobei 94 davon jeweils nur in einem Review und 13 Instrumente in zwei verschiedenen Reviews angegeben wurden. Sechs Instrumente zur Bewertung von GI (EQIP 36 (26), LIDA Instrument (22), SAM (27), brief DISCERN (28), MedlinePlus Health Information (29), AMA (30)) kamen jeweils in drei verschiedenen Reviews zur Sprache. Sechs weitere Instrumente wurden in vier Reviews genannt. Fünf davon behandeln die Lesbarkeit von GI. Diese sind der Flesch Rading Ease Score (FRES) (20), der Flesch Kinkaid Reading Grade (FKRG) (18), SMOG (23) und der Gunning Fog Index (31). Das sechste ist die Richtlinie von Silberg (32). DISCERN (16) und HON (19) wurden in zehn verschiedenen Reviews angesprochen und wurden somit mit Abstand am häufigsten genannt.

Anzahl Instrumente	Häufigkeit Nennung der Instrumente in verschiedenen Reviews
2	10
6	4
6	3
13	2
94	1
Gesamt: 121	

Tabelle 3 Häufigkeit Instrumente in den Reviews

3.3 Ergebnisse Originalpublikationen

Bei der systematischen Literaturrecherche mit anschließendem Abstract- und Volltextscreening wurden acht Originalpublikationen identifiziert. Zuvor wurden im Volltextscreening 15 Volltexte ausgeschlossen. Am häufigsten kam das Ausschlusskriterium zu tragen, dass es sich nicht um ein Manual oder um ein Instrument für die Erstellung oder Bewertung von GI und/ oder Entscheidungshilfen gehandelt hat. Danach folgte das Ausschlusskriterium, dass man das Instrument nur für GI aus einem spezifischen Themenfeld nutzen kann.

Vier verschiedene Instrumente (EQIP(26), DISCERN(16), IPDASi(33), PEMAT(34)) wurden in den acht inkludierten Originalpublikationen behandelt. Hierbei ist bei drei

Instrumenten der Fall, dass ihre Entwicklung von zwei oder sogar von drei Originalpublikationen dargelegt wurde, da es sich dabei um eine Abwandlung oder um Weiterentwicklungen handelt. Für das Instrument EQIP liegt zum einen die Version für das EQIP 20 (35) und zum anderen für EQIP 36 (16) vor. Bei EQIP 36 handelt es sich um eine Weiterentwicklung von EQIP 20, bei der 16 weitere Kriterien zur Bewertung von GI hinzugefügt wurden (26). Für das DISCERN-Instrument (16) liegt zum einen die Originalpublikation zur grundlegenden Version, zum anderen auch der Volltext für das brief DISCERN (28) – die Kurzversion von DISCERN – vor, welche zehn Jahre später erst verfasst wurde. Dabei wurden die 16 Items in der Normalversion auf zwei Faktoren und sechs Items reduziert. Drei der inkludierten Originalpublikationen behandeln das Instrument zur Bewertung einer Entscheidungshilfe für Patienten IPDASi. Im ersten Volltext, welcher 2006 erschienen ist, haben Elwyn et al. eine Kriteriensammlung für IPDASi entwickelt (36). 2009 wurden dann in einer weiteren Originalpublikation von Elwyn et al. mit dieser Kriteriensammlung IPDASi v3.0 – eine Entscheidungshilfe für Patienten mit 47 Items – entwickelt (37). Zusätzlich wurde in diesem Volltext eine Kurzversion mit 19 Items, das sogenannte IPDASi-SF, erstellt. Die dritte Originalpublikation ist eine Weiterentwicklung des IPDASi v3.0 zu IPDASi v4.0, welches aus 44 Items in drei neuen Kategorien besteht (33). Die letzte inkludierte Originalpublikation behandelt das Instrument PEMAT (34). Von diesem Instrument wurden zwei verschiedene Versionen entwickelt: Die eine ist für die Bewertung von GI aus Printmedien (PEMAT-P) und die andere Version ist für audiovisuelle GI (PEMAT-A/V) gedacht.

Publikation	Autor	Jahr	Titel
IPDASi	Elwyn, G.	2006	Developing a quality criteria framework for patient decision aids: online international Delphi consensus process
IPDASi	Elwyn, G.	2009	Assessing the quality of decision support technologies using the International Patient Decision Aid Standards instrument (IPDASi)
IPDASi4.0	Joseph-William, N.	2014	Toward Minimum Standards for Certifying Patient Decision Aids: A Modified Delphi Consensus Process

EQIP36	Charvet-Berard, A	2008	Measuring quality of patient information documents with an expanded EQIP scale
EQIP20	Moult, B.	2004	Ensuring quality information for patients: development and preliminary validation of a new instrument to improve the quality of written health care information
PEMAT	Shedlosky-Shoemaker, R.	2014	Development of the Patient Education Materials Assessment Tool (PEMAT): A new measure of understandability and actionability for print and audiovisual patient information
DISCERN	Charnock, D.	1999	DISCERN: an instrument for judging the quality of written consumer health information on treatment choices
brief discern	Khazaal, Y.	2009	Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites

Table 4 Ergebnisse Originalpublikationen

3.4 Beschreibung der einzelnen Instrumente

3.4.1 DISCERN

Das Instrument DISCERN (16) wurde 1999 von D. Charnock für Patientinnen/ Patienten und Informationsanbieterinnen/ Informationsanbieter als Zielgruppe entwickelt. Das Instrument soll bei geschriebenen Materialien - die über Therapieentscheidungen informieren - angewendet werden. Der Fragebogen besteht dabei aus 16 Items: Die ersten acht Items drehen sich um die Frage, ob die Publikation zuverlässig ist. Weitere sieben Fragen behandeln die Qualität der Information über Therapieentscheidungen. Für die ersten fünfzehn Items gibt es anschließend für die Benutzerin/ den Benutzer genaue Hinweise, worauf die Benutzerin/ Benutzer im Text besonders achten muss - wie zum Beispiel in Form von Textphrasen - um die Fragen möglichst präzise beantworten zu können. Das letzte Item ist eine Gesamtbewertung der Publikation. Dabei soll basierend auf den Antworten der obigen Fragen die Gesamtqualität der Publikation als Informationsquelle über Therapieentscheidungen bewertet werden. Bei jedem Item gilt eine Bewertungsskala von

eins bis fünf. Eins steht dabei für ein nicht erfülltes Kriterium bei der Publikation und fünf dafür, dass ein Kriterium bei der Publikation erfüllt wurde. Daraus entsteht eine Gesamtrange von 16 bis 80.

Die Fragen bezüglich der Zuverlässigkeit der Publikation drehen sich zum einen darum, ob die Ziele der Publikation klar definiert sind, die Ziele erreicht wurden und ob die Ziele relevant sind. Zum anderen wird überprüft, welche Informationsquellen für die Publikation genutzt wurden und wann die Informationen erstellt wurden. Eine weitere Frage ist die Ausgewogenheit und die Unvoreingenommenheit der Publikation. Die letzten zwei Items handeln davon, ob es Einzelheiten über zusätzliche Quellen zur Unterstützung und für weitere Informationen enthält und, ob die Publikation Angaben zu Unsicherheiten macht.

Bei den Items, welche die Qualität der Information bezüglich Therapieentscheidungen behandelt, wird einmal gefragt, ob beschrieben ist, wie jede Therapie funktioniert. Außerdem wird geprüft, ob der Nutzen und die Risiken jeder Therapie beschrieben sind und ob beschrieben wird, was passiert, wenn keine Therapie angewendet wird. Weitere Frage sind, ob die Auswirkung der Therapieentscheidung auf die Gesamtlebensqualität beschrieben wird, ob es klar ist, dass es mehr als eine mögliche Therapieentscheidung gibt und ob sie Unterstützung für die gemeinsame Entscheidungsfindung bietet.

Zehn Jahre später wurde von Y. Khazaal ein Kurzversion des DISCERN-Instruments – das brief DISCERN (28) – entwickelt. Die Kurzversion wurde auf Gesundheitswebsites, die sich mit psychischen Erkrankungen beschäftigen, angewendet und enthält sechs Items. In dieser Version gibt es noch zwei Items, welche die Zuverlässigkeit der Publikation betreffen. Das sind die Fragen bezüglich der Transparenz der Zusammenstellung der Informationsquellen und der Zeitpunkt der Informationszusammenstellung entscheidend. Vier Items beziehen sich auf die Qualität der Therapieentscheidungsinformation. Für die Überprüfung der Qualität ist einerseits entscheidend, ob beschrieben wird, wie jede Therapie funktioniert und ob der Nutzen und die Risiken jeder Therapie beschrieben wurde und andererseits, ob die Auswirkung der Therapieentscheidung auf die Gesamtlebensqualität beschrieben wurde. Das brief DISCERN hat eine Gesamtrange von 6 bis 30, wobei man bei einem Score ≥ 16 stark von der inhaltlichen Qualität überzeugt sein kann.

3.4.2 EQIP 36

Das Instrument EQIP 36 (26) wurde 2008 von A.I. Charvet-Berard entwickelt. Das Instrument kann bei Informationsdokumenten für Patientinnen/ Patienten angewendet werden. Dabei handelt es sich um 36 Fragen, die entweder mit „ja“, „teilweise“, „nein“ oder

„trifft nicht zu“ beantwortet werden können. Durch eine Formel, die die Summe der jeweiligen Antwortkategorien beinhaltet, ergibt sich ein Score, durch den man auf die Qualität von Gesundheitsinformationen Rückschlüsse ziehen kann. EQIP 36 ist eine Weiterentwicklung des Instruments EQIP 20 (35), welches 2004 von B. Moulton entwickelt wurde und wie der Name sagt, aus 20 Kriterien bestand.

Die ursprünglichen 20 Fragen können entweder mit „ja“, „teilweise“, „nein“ oder „nicht anwendbar“ beantwortet werden. Dabei zählen „ja“-Antworten als ein Punkt, „teilweise“-Antworten als ein halber Punkt und „nein“-Antworten als null Punkte. Der endgültige Score ergibt sich aus dem Quotient der Summe der einzelnen Antworten und der Differenz von 20 und der Anzahl der nicht anwendbaren Items.

Das erste Item handelt davon, ob die Publikation klar definierte Ziele hat und diese auch erreicht. Die nächsten Items betreffen die alltagsgebräuchliche Sprache, die Verwendung von kurzen Sätzen, die direkte Ansprache des Lesers und der Gebrauch einer respektvollen Schreibweise. Des Weiteren geht es um die Gestaltung der Informationen, ob es leicht verständliche Illustrationen, Diagramme oder Fotos gibt, alles in einer logischen Ordnung präsentiert wird und ob es Platz für Notizen gibt. Die nächste Frage ist, ob Kontaktdetails für Gesundheitsdienste genannt wurden. Drei weitere Items fragen, ob beschrieben ist, wann und von wem die Informationen und ob diese Informationen mit Hilfe von Dienstleistungsnutzerinnen/ Dienstleistungsnutzer erstellt wurden. Ob Hinweise auf Fragen der Lebensqualität enthalten sind, ist ein weiteres Item. Zudem wird die Verwendung von generischen Namen für Medikamente überprüft und ob Markennamen als solche gekennzeichnet sind. Ein weiteres Item ist, ob es Details von anderen Informationsquellen enthält. Die letzten Fragen befassen sich damit, ob eine Beschreibung des Zwecks, des Benefits, der Risiken und Nebenwirkungen und der Alternativen vorhanden ist.

Beim EQIP 36 sind anders als beim EQIP 20 die Fragen in drei Kategorien unterteilt. In der ersten Kategorie geht es um die inhaltliche Komponente der GI, welche mit 18 Fragen überprüft wird. Zum einen geht es um eine initiale Definition, welche Themen beinhaltet sind, und anschließend darum, ob die Themen wirklich abgedeckt sind. Des Weiteren wird abgefragt, ob eine Beschreibung des medizinischen Problems, eine Definition des Zwecks der medizinischen Intervention und eine Auseinandersetzung mit Fragen zur Lebensqualität vorliegt. Außerdem wird die GI geprüft, ob sie eine Beschreibung von Therapiealternativen, des Ablaufs der medizinischen Verfahren, des qualitativen und quantitativen Nutzens, der qualitativen und quantitativen Risiken und Nebenwirkungen, Erwähnung von Warnzeichen und von zu ergreifenden Vorsichtsmaßnahmen beinhaltet. Zusätzlich wird im EQIP 36

darauf eingegangen, ob eine Beschreibung wie mit potenziellen Komplikationen umgegangen wird und eine Auseinandersetzung mit Kosten der medizinischen Intervention und Versicherungsfragen vorhanden ist. Zusätzlich wird abgeklärt, ob spezifische Kontaktangaben für einen Krankenhausservice, sowie spezifische Angaben zu anderen Quellen von zuverlässiger Information und Unterstützung, enthalten sind. Das letzte Kriterium ist ein Gesamtbeurteilungskriterium für alle Fragen, die die inhaltliche Komponente behandeln. Dabei soll nochmals beantwortet werden, ob das Dokument alle relevanten Fragen zum Thema enthält.

Die nächste Kategorie enthält sechs Kriterien bezüglich Kenndaten. Dabei wird kontrolliert, ob das Datum der Ausstellung und ein Logo der ausstellenden Behörde vorhanden sind. Des Weiteren sind die Namen der Personen oder Einrichtungen wichtig, die zum einen das Dokument erstellt haben und zum anderen es finanziert haben. Das Vorhandensein einer kurzen Bibliografie der im Dokument verwendeten evidenzbasierten Daten wird auch geprüft. Das sechste Kriterium gibt an, ob bei der Dokumentenerstellung Patientinnen/Patienten einbezogen bzw. konsultiert wurden.

In der letzten Kategorie geht es um die strukturelle Komponente, welche zwölf Kriterien beinhaltet. Dabei wird abgefragt, ob eine alltägliche Sprache mit Erklärungen von komplexen Wörtern, kurze Sätze mit weniger als fünfzehn Wörter im Durchschnitt und generische Namen für alle Medikamente oder Produkte verwendet wird. Ein weiteres Kriterium ist ein respektvoller Ton und dass das Dokument direkt an den Leser adressiert ist. Die Information soll klar sein, eine Balance zwischen Risiko und Nutzen beinhalten und eine logische Ordnung präsentieren. In zwei weiteren Kriterien geht es um das Design, Layout und Figuren oder Grafiken, die zufriedenstellend, klar und relevant sein sollen. Außerdem wird geprüft, ob es für den Leser genügend Platz für Notizen gibt und ob das Dokument eine Einverständniserklärung entgegen den Empfehlungen enthält.

3.4.3 HONcode

Der HONcode (19) wurde von der Health On the Net Foundation für Internetwebsites erstellt. Einerseits ist dieser für Inhaberinnen/ Inhaber einer Domäne entwickelt worden, damit Websites eine entsprechende hohe Qualität aufweisen. Andererseits ist dieses Instrument für Menschen entwickelt worden, die im Internet nach Informationen suchen, um ihnen das Beurteilen von der Qualität von GI zu erleichtern. Dabei ist der HONcode eine Leitlinie, die sich in acht verschiedene Themengebiete unterteilt und den qualitativen Standard und Anforderungen einer Website aufzeigt. Im ersten Bereich geht es um die klare

Angabe von Einzelheiten über das Redaktionsteam und das Team der Website. Als zweites soll überprüft werden, ob auf der Website eine Klarstellung vorhanden ist, dass die Information nicht den direkten Arzt-Patienten-Kontakt ersetzen kann, sondern dies nur eine Ergänzung ist. Das nächste Themengebiet sind die Datenschutzrichtlinien: Diese müssen leicht zugänglich sein. Im vierten Abschnitt geht es darum, dass das Datum der letzten Aktualisierung der Website, von rechtlicher und medizinischer Seite angegeben ist und das zusätzlich auch die Quellen der Gesundheitsinformationen angegeben sind. Der fünfte Themenbereich handelt von der Präsentation der Gesundheitsinformationen. Informationen sollen vollständig und in einer objektiven, ausgewogenen und transparenten Weise zur Verfügung gestellt werden. Dies bedeutet, dass für alle Therapien, Medikamente und Operationen die jeweiligen Kontraindikationen, Nebenwirkungen, Interaktionen und Vorsichtsmaßnahmen für den Gebrauch angegeben sein müssen. Außerdem müssen Firmennamen klar zu identifizieren sein und ebenfalls sollen die Gesundheitsinformationen immer einen Hinweis enthalten, dass auch andere Firmen existieren. Der nächste Punkt behandelt die Transparenz der Website. Dabei soll die Handhabung der Website einfach, der Auftrag klar und Informationen zum Team leicht zugänglich sein. Die finanzielle Offenlegung in Zusammenhang mit einer Transparenz und klaren Angabe von allen Finanzierungsquellen ist ein weiterer wichtiger Punkt. Im letzten Abschnitt wird die Werbepolitik behandelt. Dabei sollen alle Anzeigen angegeben sein und vom Inhalt klar unterscheidbar sein.

3.4.4 IPDASi v4.0

Das Instrument IPDASi bewertet Patientenentscheidungshilfen hinsichtlich ihres Entwicklungsprozesses und Gestaltung der gemeinsamen Entscheidungsfindungskomponenten. Erstmals arbeitete G. Elwyn 2006 Kriterien zur Bewertung von Patientenentscheidungshilfen aus (14). Drei Jahre später wurde in vier Stadien daraus das Instrument IPDASi v3.0 (37) erarbeitet, welches 47 Items in zehn verschiedenen Kategorien enthält. Außerdem wurde das IPDASi-SF – eine Kurzversion mit 19 Items – entwickelt. Aus IPDASi v3.0 wurde durch eine Weiterentwicklung von N. Joseph-Williams im Jahre 2013 IPDASi v4.0 (33) erstellt. Dabei wurde die Anzahl der Items auf 44 reduziert.

Grundsätzlich werden bei IPDASi v4.0 alle Items in drei verschiedene Kategorien eingeteilt, welche die Wichtigkeit der einzelnen Items darstellen. Zum einen gibt es die qualifizierenden Kriterien: Diese sind auf jeden Fall erforderlich, um eine Intervention als

Entscheidungshilfe ansehen zu können. Davon gibt es bei IPDASi 4.0 insgesamt sechs. Des Weiteren gibt es Zertifizierungskriterien: Ohne diese Kriterien hat eine Entscheidungshilfe ein hohes Risiko eine schädliche Verzerrung aufzuweisen. Das Instrument IPDASi v4.0. enthält hiervon zehn Items. Die letzte Gruppe beinhaltet insgesamt 28 Kriterien und ist somit die größte Kategorie: Dies ist die Gruppe der Qualitätskriterien. Bei diesen Kriterien wird angenommen, dass sie zwar eine Entscheidungshilfe stärken, aber ihre Abwesenheit kein hohes Risiko einer schädlichen Verzerrung darstellt.

Die 44 Kriterien werden außerdem in zehn verschiedene inhaltliche Kategorien untergeordnet. Die Kategorien betreffen einmal die Information, die Wahrscheinlichkeit, den Wert, die Führung, die Entwicklung, die Nachweise, die Veröffentlichung, die klare Sprache, die Bewertung und die Tests. Aufgrund der hohen Wichtigkeit werden anschließend die qualifizierenden Kriterien und die Zertifizierungskriterien des Instruments genauer beschrieben und von der Beschreibung der einzelnen Qualitätskriterien aufgrund ihrer hohen Anzahl abgesehen. Die Kategorien bezüglich der Wahrscheinlichkeit, der Führung, der Entwicklung, der klaren Sprache und der Bewertung beinhalten nur Qualitätskriterien. Dies bedeutet, dass nur fünf der zehn Kategorien Kriterien enthalten, die für die Bewertung der Qualität der Patientenentscheidungshilfe relevant sind und nicht weggelassen werden dürfen.

Fünf der sechs qualifizierenden Kriterien sind Items, welche Informationen über die einzelnen Optionen der Entscheidungsfindung abfragen. Das erste Item fordert eine Beschreibung des Gesundheitszustandes oder des Problems der Patientenentscheidungshilfe für das die Entscheidung erforderlich ist. Ein weiteres Kriterium ist eine explizite Angabe der Entscheidung, die zu berücksichtigen ist. Des Weiteren wird nach einer genauen Beschreibung der verfügbaren Optionen gefragt. Die Beschreibung der positiven und negativen Aspekte jeder Option in der Patientenentscheidungshilfe sind zwei weitere Items. Das sechste qualifizierende Kriterium ist, ob die Patientenentscheidungshilfe beschreibt, wie sich die Konsequenzen der Optionen genau anfühlen. Dabei sind die physischen, psychischen und sozialen Konsequenzen gemeint.

Bezüglich der Information einer Patientenentscheidungshilfe gibt es zusätzlich zu den fünf qualifizierenden Kriterien noch ein Zertifizierungskriterium. Bei diesem Item geht es darum, ob die Patientenentscheidungshilfe die negativen und positiven Eigenschaften der Optionen in gleicher Ausführlichkeit angibt. Des Weiteren betreffen vier Zertifizierungskriterien die Nachweise einer Patientenentscheidungshilfe: Dabei ist wichtig, ob eine Patientenentscheidungshilfe Zitate zu den ausgewählten Evidenzen und ein Herstellungs-

und Veröffentlichungsdatum liefert, Auskunft über die Aktualisierungsrichtlinie gibt und ob sie über den Grad der Unsicherheit über Ereignis- oder Ergebniswahrscheinlichkeiten informiert. Die Veröffentlichung betreffend gibt es ein Item, ob die Patientenentscheidungshilfe Informationen über die Finanzierungsquellen der Entwicklung liefert. Schlussendlich existieren noch vier Zertifizierungskriterien, welche grundlegende Tests betreffen: Dabei ist entscheidend, ob beschrieben wird, was im jeweiligen Test genau gemessen werden soll und ob die Patientenentscheidungshilfe jeweils die nächsten typischen Schritte, die gemacht werden sollen, angibt, falls der Test das Problem oder den Zustand entdeckt bzw. nicht entdeckt. Das letzte Zertifizierungskriterium ist, ob die Patientenentscheidungshilfe Informationen über die Konsequenz der Zustands- oder Krankheitsentdeckung hat, die nie Probleme verursacht hätte, wenn das jeweilige Screening nicht durchgeführt worden wäre.

3.4.5 PEMAT

Das Instrument PEMAT wurde 2014 von S. Shoemaker et al. in einer Originalpublikation veröffentlicht (34). Es kann Fachleute bei der Bewertung der Qualität von Gesundheitsinformationen unterstützen. Grundsätzlich fasst das Instrument 26 Items und zwei Dimensionen – Verständlichkeit (19 Items) und Handlungsfähigkeit (sieben Items). Außerdem beinhaltet das Instrument PEMAT zwei verschiedene Versionen: Erstens enthält es eine Version für Printmaterial (PEMAT-P). Dieses enthält 17 Items bezüglich der Verständlichkeit und sieben Items in der Kategorie Handlungsfähigkeit. Und Zweites gibt es noch eine Version für audiovisuelles Material (PEMAT-A/V) und diese Version hat dreizehn Items in der Kategorie der Verständlichkeit und vier bezüglich der Handlungsfähigkeit. Das bedeutet, dass manche Items bei Printmaterial und bei audiovisuellem Material zutreffen und manche nur bei einem von beidem. Im folgenden Text werden erst die Items in der Dimension der Verständlichkeit behandelt: Zwölf der 19 Items können für Printmaterial und audiovisuelles Material angewendet werden. Das erste Item ist, ob das Material seinen Zweck vollständig ersichtlich macht. Im Weiteren wird behandelt, ob das Material gewöhnliche und alltägliche Sprache verwendet, ob die aktive Ansprache benutzt wird und ob medizinische Begriffe genau definiert werden. Außerdem soll das Material in kurze Abschnitte mit informativen Überschriften unterteilt sein, eine logische Reihenfolge haben und optimalerweise eine Zusammenfassung haben. Ein weiteres Item fragt nach visuellen Hinweisen, wie zum Beispiel Pfeile, Hervorhebungen und Aufzählungszeichen, die dabei helfen sollen, die Aufmerksamkeit auf die wichtigen Dinge

zu lenken. Außerdem wird gefragt, ob das Material einerseits übersichtliche und klare Illustrationen und Fotos und andererseits einfache Tabellen mit kurzen und klaren Zeilen- und Spaltenüberschriften verwendet.

Bei Printmaterialien gibt es zusätzliche fünf Items. Diese behandeln, ob das Material keine Informationen oder Inhalte, die von ihrem Zweck ablenken beinhalten, im Material vorkommende Zahlen klar zu verstehen sind und vom Benutzer keine Durchführung von Berechnungen zu erwarten sind. Außerdem handeln zwei Items von den visuellen Hilfsmitteln des Materials. Diese sollen zum einen den Inhalt verstärken und nicht ablenken und sollen klare Titel oder Beschriftungen haben. Für audiovisuelle Materialien gibt es noch zwei andere Items: Der Text auf dem Bildschirm soll einerseits einfach zu lesen sein und der Benutzer soll die Wörter klar hören können und somit nicht zu schnell oder undeutlich sein. Nun werden die Items bezüglich der Handlungsfähigkeit behandelt: Für Printmaterial und audiovisuelles Material können vier der sieben Items angewendet werden. Im ersten Item geht es um darum, ob das Material mindestens eine Aktion eindeutig identifiziert, die der Benutzer ausführen kann. Außerdem wird gefragt, ob bei der Beschreibung von Aktionen der Leser direkt angesprochen wird und ob das Material jede Handlung in überschaubare Schritte gliedert. Das letzte Item bezüglich der Handlungsfähigkeit, welches Printmaterialien und audiovisuelle Materialien betrifft, handelt, ob Erklärungen existieren, wie zum Beispiel die Tabellen, Diagramme und Schaubilder zur Entscheidungsfindung verwendet werden können. Es gibt kein weiteres Item für audiovisuelle Materialien, aber drei zusätzliche für Printmaterialien. Zum einen wird nach greifbaren Werkzeugen gefragt, die dem Benutzer es erleichtern sollen, eine Maßnahme auszuführen. Zum anderen, ob das Material visuelle Hilfsmittel verwenden, welche zusätzlich die Anweisungsbefolgung erleichtern können. Außerdem wird geprüft, ob das Material das Angebot für Anweisungen oder Beispiele für Berechnungen bietet.

3.4.6 LIDA

Das LIDA Instrument (22) wurde von dem Unternehmen Minervation herausgegeben, dessen Gründer André Tomlin und Douglas Badenoch sind. Dies ist ein Instrument zur Validierung von Gesundheitswebsites und soll bei der Erstellung von nützlichen und zuverlässigen Gesundheitsinformationen – vor allem Gesundheitswebsites - helfen. Allgemein ist das Instrument in drei Bereiche unterteilt: Zugänglichkeit, Verwendbarkeit und Zuverlässigkeit. Es wird eine Skala von null bis drei verwendet, bei der null „niemals“ entspricht und drei „immer“. Insgesamt gibt es 53 Items und es können 159 Punkte erreicht

werden. 63 Punkte können maximal im Bereich der Zugänglichkeit der Gesundheitsinformation erreicht werden. Dabei wird die Gesundheitswebsite zum Beispiel bezüglich Seiteneinstellungen, wie zum Beispiel HTML Sprachdefinition, Meta Tag Schlüsselwörter und Seitentitel oder Zugangsbeschränkungen und Dublin Core-Titel-Tags überprüft. Die Items bezüglich der Verwendbarkeit werden noch mal in vier Untergruppen unterteilt: Klarheit, Konsistenz, Funktionalität und Erreichbarkeit. Die Untergruppe „Klarheit“ beinhaltet dabei sechs Fragen. Diese prüfen beispielsweise, ob eine klare Aussage und eine gewisse Übersichtlichkeit mittels einer guten Navigation und Layout vorhanden sind. Hinsichtlich der Konsistenz wird das Seitenlayout, die Navigation und die Struktur der Seite überprüft. Die Untergruppe „Funktionalität“ enthält fünf Items, die sich mit der Handhabung der Website auseinandersetzen. Zum Beispiel bezieht sich ein Item auf die Frage, ob eine effektive Suchfunktion vorhanden ist. Ein anderes fragt, ob die Website auch ohne Plug-ins von Drittanbietern verwendet werden kann. Die letzte Untergruppe, die sich auf die Verwendbarkeit der Gesundheitswebsite bezieht, ist die Ansprechbarkeit. Ob die Benutzerinnen/ Benutzer selbstständig beurteilen kann, ob die Seite auf ihn zutrifft, ist dabei ein Item. Außerdem ist in dieser Untergruppe relevant, ob die Seite interaktiv ist und auch nicht-textuelle Medien, wie zum Beispiel Grafiken, audiovisuelle Medien und Forschungsdaten enthält. Der dritte Abschnitt des LIDA Instruments behandelt die Zuverlässigkeit der Gesundheitswebsite. Allgemein ist dieser Abschnitt in fünf Untergruppen unterteilt, wobei zwei von diesen ergänzend und somit optional sind. Im Gesamten können 51 Punkte erreicht werden und dabei sind dann 17 Items abgefragt. Die erste Untergruppe betrifft die Aktualität und fragt nach, ob die Website auf aktuelle Ereignisse reagiert, spezifische Inhalte von Benutzerinnen/ Benutzern kommentiert werden können und ob regelmäßig der Inhalt aktualisiert wird. In der nächsten Untergruppe werden mögliche Interessenskonflikte behandelt. Für den Erhalt der Maximalpunktzahl von neun Punkten muss klar sein, wer die Seite betreibt und für sie bezahlt und es muss eine Erklärung über die Ziele der Personen, die die Website betreiben, vorhanden sein. Die letzte obligatorische Untergruppe bezieht sich auf die Inhaltserstellung, welche anhand einer klaren und robusten Methode erfolgen sollte. Zusätzlich sollte die Information anhand von Originalquellen überprüft werden können. In der einen optionalen Untergruppe werden die Verfahren zur Produktion von Inhalten genauer erfragt und die andere optionale Untergruppe behandelt die Ausgabe des Inhalts der Gesundheitswebsite.

3.4.7 Suitability of Materials (SAM)

Das Instrument SAM wurde von Doak C. und Doak L. (27) zur Bewertung von GI entwickelt. Vor allem soll es bei der Beurteilung von Printmaterial, Abbildungen, Video- und Tonbandanweisungen angewendet werden. Insgesamt besteht das Instrument aus 22 Faktoren, die jeweils mit Punkten bewertet werden sollen. Dabei stehen null Punkte für nicht geeignet, ein Punkt für ausreichend und zwei Punkte für überragend. Außerdem kann die Anwenderin/ der Anwender einen Faktor auch aus „N/A“ kennzeichnen, wenn der Faktor für das Material nicht geeignet sein soll. Das endgültige Ergebnis ist eine Prozentpunktzahl, die errechnet wird, indem die addierten Punkte der einzelnen Faktoren durch die mögliche Höchstpunktzahl dividiert wird. Dabei wird ein Material mit einem Endergebnis von 70-100% als überragend eingestuft, ein Material mit 40-69% als ausreichend und Material mit 0-39% als ungeeignet. Es gibt sechs verschiedene Kategorien, in die die 22 Faktoren eingeteilt sind. Die erste Kategorie betrifft den Inhalt. Dabei wird beurteilt, ob der Zweck ersichtlich ist, der Umfang begrenzt ist, ein Inhalt über Verhaltensweisen und eine Zusammenfassung vorhanden ist. In der zweiten Kategorie wird die Anforderung an die Lesefähigkeit bewertet: Dazu gehört zum einen die Leseniveaustufe und der Schreibstil, welcher möglichst aktiv gehalten werden soll und zum anderen ein Vokabular mit möglichst gewöhnlichen Wörtern, informative Überschriften und viele präsentierte Zusammenhänge. Der nächste Punkt betrifft die Grafiken des Materials. Wichtig dabei ist zum einen die Deckblattgrafik, welche ihren Zweck zeigen soll, und zum anderen die Art, Beschriftung und Erklärung der einzelnen Grafiken und die Relevanz der Abbildungen. Im nächsten Punkt werden die Layoutfaktoren, die Typografie und das Vorhandensein von Zwischenüberschriften bewertet. Die Verwendung von Interaktionen, die Modellierung von Verhaltensweisen und die Motivation wird in der vorletzten Kategorie betrachtet. Aus zwei Faktoren besteht die letzte Kategorie. Diese befasst sich mit der kulturellen Angemessenheit und misst die Übereinstimmung in Logik, Sprache und Erfahrung und das kulturelle Bild.

3.5 Beschreibung der Instrumente zur Beurteilung der Lesbarkeit

Bei der systematischen Literaturrecherche wurde neben den Instrumenten zur Bewertung der Qualität von Gesundheitsinformationen auch einige Instrumente ausfindig gemacht, die die Lesbarkeit und Verständlichkeit der Texte beurteilen. Dabei wurden der Flesch Kinkaid Reading Grade (18), der Flesch Reading Ease Score (20), SMOG (23) und der Gunning Fog Index (31) – wie auch schon in den Ergebnissen genau aufgeschlüsselt - am häufigsten

verwendet. Alle vier Instrumente wurden nicht spezifisch für die Beurteilung von medizinischen Paper entwickelt, sondern werden von verschiedensten Fachgebieten verwendet.

3.5.1 Flesch Reading Ease Score (FRES)

1948 entwickelte Rudolf Flesch den Flesch Reading Ease Score (20). Dieser besteht aus zwei Komponenten: Zum einen wird die Anzahl der Wörter im Text zu der Anzahl der Sätze im Text im Verhältnis genommen und daraus ergibt sich dann die durchschnittliche Satzlänge (ASL). Und zum anderen wird der Quotient von der Silbenanzahl des gesamten Textes und der Wörteranzahl des Textes berechnet (ASW).

$$FRES = 206,835 - (1,015 \cdot ASL) - (58,5 \cdot ASW)$$

Je höher der Wert des FRES ist, desto verständlicher ist der Text.

Pattern of "Reading Ease" Scores

"Reading Ease" Score	Description of Style	Typical Magazine	Syllables per 100 words	Average Sentence Length in Words
0 to 20	Very Difficult	Scientific	192 or more	29 or more
30 to 50	Difficult	Academic	167	25
50 to 60	Fairly difficult	Quality	155	21
60 to 70	Standard	Digests	147	17
70 to 80	Fairly easy	Slick-fiction	139	14
80 to 90	Easy	Pulp-fiction	131	11
90 to 100	Very easy	Comics	123 or less	8 or less

Tabelle 5 Einteilung des FRES

Da im Deutschen durchschnittlich die Wörter länger sind, aber die Sätze meist gleich lang bleiben, hat Toni Amstad (38) eine passende Version für deutsche Texte entwickelt. Dabei wurden die beiden Komponenten beibehalten und nur die Gewichtung dieser in der Formel verändert.

$$FRES_{deutsch} = 180 - ASL - (84,6 \cdot ASW)$$

3.5.2 Flesch Kinkaid Reading Grade (FKGL)

Der Flesch Kinkaid Reading Grade (FKGL) ist eine Weiterentwicklung des FRES von J. Peter Kincaid et al. (18). Es wurde für die US Navy weiterentwickelt, um für die Marine besser geeignet zu sein. Bei diesem Instrument wird die Lesbarkeit eines Textes durch die

Anzahl der Schuljahre angegeben wird, die der Leser absolviert haben muss, um den Text verstehen zu können. Aufgrund der Entwicklung des Instruments in der USA ist das Ergebnis auf das amerikanische Schulsystem abgestimmt. Dabei wurden die zwei Komponenten vom FRES unverändert übernommen und damit eine neue Formel kreiert:

$$FKGL = (0,39 \cdot ASL) + (11,8 \cdot ASW) - 15,59 \cdot$$

3.5.3 Gunning Fog Index (GFI)

Der Gunning Fog Index (GFI) wird von Robert Gunning (31) im Jahre 1968 entwickelt. Dieses Instrument gibt ebenso wie der FKGL die Anzahl der Schuljahre an, die für das Leseverständnis absolviert werden müssen. Dabei lehnt sich der GFI ebenfalls an das amerikanische Schulsystem. Für die Berechnung wird eine Textstelle gewählt, die zumindest eine Länge von 100 Wörtern hat. Anschließend wird erneut - wie auch beim FRES und FKGL - die Anzahl der Wörter im Verhältnis zu der Anzahl der Sätze genommen (ASL). Im zweiten Schritt werden alle Wörter mit drei oder mehr Silben gezählt (D) und diese im Verhältnis zu der Gesamtanzahl der Wörter (W) gesehen.

$$GFI = \left(ASL + 100 \frac{D}{W} \right) \cdot 0,4$$

3.5.4 Simple Measure of Gobbledygook for readability (SMOG)

Im Jahre 1969 hat G. Harry McLaughlin SMOG (23) entwickelt. Dies ist eine Weiterentwicklung des GFI. Bei diesem Instrument werden 30 Sätze gezählt: Dabei sollen etwa 10 Sätze vom Anfang, 10 Sätze von der Mitte und die letzten 10 Sätze vom Ende sein. Danach werden alle mehrsilbigen Wörter – also Wörter mit drei oder mehr Silben – gezählt (NOP). Wenn gleiche Wörter öfters vorkommen, werden sie auch mehrfach gezählt. Diese Zahl wird daraufhin in folgende Formel eingesetzt:

$$grade = 1,0430\sqrt{NOP} + 3,1291$$

In einer Studie von Conteras, A. et al (39) erwies sich das Instrument SMOG bei verbrauchsorientierten Gesundheitsmaterialien im Vorteil zum FKRG. Dabei zeigt die Studie auf, dass SMOG die Leseschwierigkeit präziser einschätzt und das Instrument FKRG diese eher unterschätzt.

3.6 Gegenüberstellung der Instrumente/ Kriterien

Im nächsten Schritt werden die Instrumente zur Bewertung der Qualität von GI miteinander verglichen. Dafür wurden die Kriterien der einzelnen Instrumente in einer Tabelle (siehe

Tabelle 6) gegenübergestellt und extrahiert, welche Kriterien wie oft genau in den verschiedenen Instrumenten verwendet wurden, um somit eine Priorisierung einzelner Kriterien aufstellen zu können. Bei dieser Analyse wurden die sieben Instrumente betrachtet, welche schon in Kapitel 3.4 genauer beschrieben wurden. Es wurde dabei immer jeweils die aktuellste und lange Version eines Instruments berücksichtigt: DISCERN (16), EQIP36 (26), IPDASi4.0 (33), HON (19), PEMAT (34), LIDA (22) und SAM (27). Da die verschiedenen Instrumente ihre Kriterien auf unterschiedliche Art und Weise darstellen und da die Instrumente unterschiedliches Gesundheitsmaterial ansprechen, wurden manche Kriterien, die prinzipiell gleiche Inhalte betrachten zusammengefasst. Kein Kriterium wurde dabei in allen sieben Instrumente verwendet. Des Weiteren ergab die Analyse der Häufigkeiten der einzelnen Kriterien, dass ein Kriterium jeweils in sechs bzw. fünf Instrumenten verwendet wurden, fünf Kriterien in vier, sechs Kriterien in drei und 16 Kriterien in zwei Instrumenten benutzt wurden.

3.6.1 In sechs Instrumenten verwendetes Kriterium

Das Kriterium einer klaren Zieldefinition wurde in ganzen sechs Instrumenten (DISCERN (16), EQIP36 (26), PEMAT (34), LIDA (22), SAM (27)) verwendet. Dabei wurde in EQIP36 nach dem Zweck der medizinischen Intervention gefragt.

3.6.2 In fünf Instrumenten verwendetes Kriterium

Dieses Kriterium beschäftigt sich mit den Quellen von Gesundheitsinformationen: DISCERN (16), LIDA (22), HON (19) und IPDASi4.0 (33) fragen ab, ob Quellenangaben vorhanden sind und ob einzelne Informationen mittels Originalquellen bzw. Evidenzen überprüft werden können. In EQIP36 ist überdies eine kurze Bibliografie der Quellen erwünscht.

3.6.3 In vier Instrumenten verwendete Kriterien

Fünf verschiedene Kriterien werden in vier Instrumenten verwendet. Die ersten beiden Kriterien beschäftigen sich damit, ob der Benefit jeder einzelnen Therapie zum einen, als auch die Risiken und Nebenwirkungen zum anderen genau beschrieben werden. Dies ist im Unterschied zu den anderen Kriterien dieser Kategorie ein inhaltliches Thema. Dieses Kriterium kommt bei DISCERN (16), EQIP36 (26), HON (19) und IPDASi4.0 (33) zur Anwendung. Bei EQIP36 wird dabei nochmals zwischen quantitativem und qualitativen Benefit bzw. Risiko unterschieden.

Das Vorhandensein des Datums der Ausstellung der Information ist ein weiteres Item, das in vier Instrumenten genannt wird (DISCERN, EQIP36, HON, IPDASi4.0).

Die letzten beiden Kriterien in dieser Kategorie betreffen die Personen, welche einerseits die Informationen erstellt haben und andererseits diese finanziert haben. Dabei ist beim EQIP36 entscheidend, wer das Paper erstellt hat bzw. finanziert hat. Bei HON, IPDASi4.0 und LIDA bezieht sich das Item auf die genaue Benennung des Betreibers bzw. Finanzierin/ Finanzier der Website bzw. der Entwicklung.

3.6.4 In drei Instrumenten verwendete Kriterien

Die Analyse ergab sechs Kriterien, die jeweils in drei Instrumenten verwendet wurden.

Dabei betrifft das erste Item die Verwendung einer alltäglichen Sprache. Dazu zählt außerdem die Erklärung von komplexen Wörtern. Auch das zweite Kriterium betrifft den Schreibstil. Dieser soll aktiv gestaltet sein und den Leser direkt ansprechen. Diese beiden Kriterien sind in EQIP36, PEMAT und SAM enthalten.

Ein Kriterium, welches bei EQIP36, PEMAT und LIDA Anwendung findet, bezieht sich auf die Struktur der GI: Dabei stützt sich EQIP36 und PEMAT auf eine logische Reihenfolge in den jeweiligen Dokumenten. Da LIDA für die Anwendung von Websites entwickelt wurde, bezieht sich das Kriterium bei LIDA auf eine klare und gut strukturierte Navigation.

Ein weiteres Kriterium bezieht sich auf das Layout der Seite bzw. des Dokuments: Dieses wird in EQIP36, LIDA und SAM verwendet. Bei LIDA wird dabei zum einen das Layout an sich bewertet – ob es klar lesbar ist – und dabei hinaus auch, ob das gleiche Seitenlayout durchgängig beibehalten wurde. Das Instrument SAM hat bei diesem Kriterium eine objektive Herangehensweise: Durch fünf genaue Unterkriterien, wird bestimmt inwieweit das Layout zufriedenstellend ist.

EQIP36, PEMAT und SAM haben alle ein Item, welches beurteilt, ob übersichtliche Figuren, Grafiken, Illustrationen und Co verwendet werden.

Das letzte Item, welches in drei Instrumenten vorkommt, betrifft die Unvoreingenommenheit der GI: Dabei bezieht sich DISCERN auf die Publikationen die ausgewogen und unvoreingenommen sein soll, EQIP36 auf die Ausgewogenheit der Information zwischen Benefit und Risiko und HON auf jegliche medizinische Information die ausgeglichen dargestellt werden sollen.

3.6.5 In zwei Instrumenten verwendete Kriterien

Es gibt 16 Kriterien, die in zwei Instrumenten zur Bewertung der Qualität von GI verwendet werden. Davon sind drei Kriterien in DISCERN und EQIP36 vertreten. Die ersten beiden gemeinsamen Kriterien sind, ob eine Beschreibung einerseits des Therapieablaufs und andererseits der Auswirkung der Therapieentscheidung auf die Gesamtlebensqualität vorhanden ist. Außerdem haben beide Instrumente eine Gesamtbewertung der Publikation beziehungsweise der GI als einzelnes Item.

Zwei Items überschneiden sich bei EQIP36 und IPDASi4.0. Zum einen wird dort nach der Beschreibung des medizinischen Problems gefragt und zum anderen ob Therapiealternativen beschrieben sind. Bei der Frage nach dem medizinischen Problem sieht man eine klare Abgrenzung zu der Frage nach der genauen Zieldefinition, dessen Kriterium schon in Abschnitt 3.6.1 beschrieben wurde.

IPDASi4.0 und PEMAT haben zusätzlich ebenso zwei gleiche Kriterien. Dabei geht es direkt um das Thema der Problemlösung. Zum einen, ob mögliche Handlungsschritte für eine Problemlösung in einzelne, kleine Schritte unterteilt ist und zum anderen, ob explizite Entscheidungen angegeben sind, die für eine Problemlösung zu berücksichtigen sind.

Das Item, das abfragt, ob beschrieben ist, wie der Verlauf ohne eine Therapie ist, wird bei DISCERN und IPDASi4.0 benutzt.

Die Instrumente EQIP36 und HON fragen beide, ob Vorsichtsmaßnahmen beschrieben wurden, die der Patient selbst ergreifen kann.

Fünf Items haben die Instrumente PEMAT und SAM gemeinsam. Dabei beschäftigen sich zwei Items mit der Übersichtlichkeit. Einerseits wird dabei nach informativen Überschriften für Textabschnitte gefragt und andererseits nach einer vorhandenen Zusammenfassung. Die anderen drei Kriterien drehen sich um visuelle Hilfsmittel. Dabei wird abgefragt, ob diese keine unwesentlichen Details zeigen, klare Titel und Beschriftungen haben und Tabellen ebenso kurze und klare Zeilen- und Spaltenüberschriften haben.

Wie hoch das Leseverständnislevel des jeweiligen Materials ist, wird von IPDASi4.0 und SAM als Kriterium benutzt.

Zuletzt haben die Instrumente LIDA und SAM ein gemeinsames Kriterium. Dieses ist die Frage, ob die Seite interaktiv ist.

	DISCERN	EQUIP 36	HON	IPDASI 4.0	PEMAT	LJIDA	SAM
klare Zieldefinition/ Zweck medizinischer Intervention	x	x	x		x	x	6
klar, welche Informationsquellen genutzt wurden, um die Publikation zusammenzustellen/ kann Information anhand von Originalquellen überprüft werden/ kurze Biografie der im Dokument verwendeten evidenzbasierten Daten/ Angabe Quellen/ Angabe Zitate	x	x	x	x		x	5
Ist Benefit jeder Therapie beschrieben?	x	x	x	x			4
Beschreibung Risiken jeder Therapie	x	x	x	x			4
Datum Informationserstellung/ Datum der Ausstellung oder Revision	x	x	x	x			4
Name/ Qualifikation/ Einzelheiten zu Referenzen des Redaktionsteam, Team der Website, Autoren		x	x	x		x	4
Name der Personen oder Einrichtungen, die das Dokument finanziert haben		x	x	x		x	4
Verwendung alltäglicher Sprache, Erklärung komplexer Wörter oder Jargon		x			x		3
an den Leser direkt adressiert/ aktive Sprache		x			x		3
Information in logischer Ordnung/ logische Reihenfolge/ Ist die Navigation klar und gut strukturiert?		x			x	x	3
Design und Layout zufriedenstellend/ Ist das Layout des Hauptblocks der Information klar und lesbar?		x				x	3
Figuren, Illustrationen, Fotos oder Grafiken sind klar, übersichtlich und relevant		x			x		3
Ist Publikation ausgewogen und unvoreingenommen/ Info ausgeglichen zw. Risiko und Benefit	x	x	x				3
Ist beschrieben, wie Therapie funktioniert?/ Ablauf Therapie	x	x					2
Beschreibung was passiert, wenn keine Therapie angewendet wird/ natürlicher Verlauf ohne Therapie	x			x			2
Wird Auswirkung der Therapieentscheidung auf die Gesamtlebensqualität beschrieben	x	x					2
Gesamtbewertung Publikation	x	x					2
Beschreibung des medizinischen Problem		x		x			2
Beschreibung Therapiealternativen		x		x			2
Beschreibung von Vorsichtsmaßnahmen, die Pat. ergreifen kann		x	x				2
Gibt explizit Entscheidungen an, die zur Problemlösungen zu berücksichtigen sind		x		x			2
bietet Schritt-für Schritt-Weg, um Entscheidung zu treffen/ gliedert jede Handlung in überschaubare, explizite Schritte				x	x		2
berichtet über Lesbarkeitslevel/ (niedriges) Leseverständnislevel				x			2
Abschnitte mit informativen Überschriften					x		2
Material bietet Zusammenfassung/ Zusammenfassung oder Übersicht enthalten					x		2
visuelle Hilfsmittel des Materials verstärken den Inhalt eher als sie vom Inhalt ablenken/ keine unwesentlichen Details					x		2
visuelle Hilfsmittel haben klare Titel oder Beschriftungen					x		2
Material verwendet einfach Tabellen mit kurzen und klaren Zeilen- und Spaltenüberschriften					x		2
Ist Seite interaktiv?						x	2
Kann Benutzer sehen, ob Website für sie zutrifft?						x	1
Kann der Benutzer seine Erfahrungen bei der Nutzung der Website personalisieren?						x	1
Integriert die Website nicht-textuelle Medien?						x	1
Reagiert die Seite auf aktuelle Ereignisse?						x	1
Können Benutzer spezifischen Inhalte kommentieren?						x	1

Tabelle 6.4 Bewertung Qualitätskriterien

	DISCERN	EQUIP 36	HON	IPDASI 4.0	PEMAT	LIDA	SAM
Wird der Inhalt der Seite in einem angemessenen Intervall aktualisiert?						x	1
Gibt es eine Erklärung über die Ziele der Personen, die die Website betreiben?						x	1
Weist die Seite eine klare Methode der Inhaltserstellung auf?						x	1
Ist die Methode robust?						x	1
Werden die Bedürfnisse des Publikums im Voraus ermittelt?						x	1
Wird eine umfassende Literaturrecherche durchgeführt?						x	1
Werden die gefundenen Dokumente kritisch betrachtet?						x	1
Würde der Inhalt von Fachexperten verfasst?						x	1
Wird der Inhalt von einem unabhängigen Experten oder Gremium überprüft?						x	1
Bietet die Seite genaue und zuverlässige Informationen?						x	1
Hat die Literatursuche die richtigen Informationen gefunden?						x	1
Stimmt der Inhalt mit der Suche überein?						x	1
Ist der Inhalt korrekt?						x	1
Inhalt über Verhaltensweisen							x
Umfang ist begrenzt							x
Zuerst Kontext angeben							x
Deckblattgrafik zeigt Zweck							x
klare Typographie (gleiche Schriftgröße, nicht durchgängige Großschreibung etc.)							x
Zwischenüberschriften/ Untertitel verwendet							x
spezifische, praktische Verhaltensweisen modelliert							x
Motivation, Selbstwirksamkeit							x
Übereinstimmung mit Logik, Erfahrung und Sprache							x
kulturelle Bilder und Beispiele in realistischer Art und Weise							x
Ziele erreicht	x						1
beinhaltet keine Informationen oder Inhalte, die von Zweck ablenken					x		1
Publikation relevant?							1
erhält Einzelheiten über zusätzliche Quellen von Unterstützung und Information	x						1
Bezieht sich auf Bereiche der Unsicherheit	x						1
beschreibt Benefit und Risiken jeder Option in jeder Ausführlichkeit				x			1
beschreibt Interaktionen der Therapien			x				1
ist klar, dass es mehr als eine mögliche Therapieentscheidung gibt?	x						1
bietet sie Unterstützung für die gemeinsame Entscheidungsfindung?	x						1
Abdeckung oben genannte Themen			x				1
Erwähnung von Warnzeichen, die Pat. erkennen kann			x				1
Auseinandersetzung mit Kosten der medizinischen Intervention und Versicherungsfragen			x				1
spezif. Kontaktangaben für Krankenhausservice			x				1
spezif. Angaben zu anderen Quellen von zuverlässiger Information/ Unterstützung			x				1
Beschreibung, wie mit potenziellen Konflikten umgegangen wird			x				1

Tabellenführung Qualitätskriterien

	DISCERN	EQIP 36	HON	IPDASI 4.0	PEMAT	LJIDA	SAM
beschreibt, wie sich Konsequenzen der Optionen anfühlen				x			1
bittet Pat. zu überlegen, welche pos. und neg. Aspekte der Optionen für sie/ihn am wichtigsten sind				x			1
Logo der ausstellenden Behörde		x					1
Dokument gibt an, ob und wie Pat. bei Erstellung der GI einbezogen wurden		x					1
Verwendung von generischen Namen für alle Medikamente oder Produkte		x					1
Verwendung kurzer Sätze		x					1
Ton ist respektvoll		x					1
Info ist klar (keine Zweideutigkeit etc.)		x					1
Dokument hat Platz für Notizen der Leser/Lesers		x					1
Dokument enthält Einverständniserklärung entgegen den Empfehlungen		x					1
Klarstellung, dass Information nicht direkten Arzt-Patienten-Kontakt ersetzen kann			x				1
Zugang zu Datenschutzrichtlinien			x				1
Werbepolitik: alle Anzeigen sollen angegeben sein und vom Inhalt unterscheidbar			x				1
Firmennamen müssen klar zu identifizieren sein			x				1
liefert Information über Ergebniswahrscheinlichkeiten, die mit Optionen verbunden sind				x			1
benennt definierte Patientinnen/Patientengruppen für die die Ergebniswahrscheinlichkeiten gelten				x			1
benennt Ereignisraten für die Ergebniswahrscheinlichkeiten				x			1
ermöglicht Vergleich v. Ergebniswahrscheinlichkeiten zw. Optionen unter Benutzung desselben Zeitraums				x			1
ermöglicht Vergleich v. Ergebniswahrscheinlichkeiten zw. Optionen unter Benutzung desselben Nenners				x			1
Bietet mehr als einen Weg die Wahrscheinlichkeiten darzustellen				x			1
enthält Hilfsmittel (Arbeitsblätter, Listen) für Diskussion der Optionen mit Ärztin/Arzt				x			1
enthält Bedarfsanalyse mit Kunden oder Pat.				x			1
Entwicklungsprozess enthält Bedarfsanalyse mit Angehörigen der Gesundheitsberufe				x			1
Entwicklungsprozess beinhaltet Überprüfung durch Pat. die nicht an der Erstellung der entscheidungsunterstützenden Intervention beteiligt waren				x			1
Entwicklungsprozess beinhaltet die Überprüfung durch Fachleute, die nicht an der Erstellung der Entscheidungshilfe beteiligt waren				x			1
wurde im Fokustest mit Pat. getestet, die mit der Entscheidung konfrontiert waren				x			1
wurde im Fokustest mit Ärzten getestet, die Patienten beraten, die vor der Entscheidung stehen				x			1
Auskunft über Aktualisierungsrichtlinie				x			1
liefert Information über Grad der Unsicherheit über Ereignis- oder Ergebniswahrscheinlichkeiten				x			1
beschreibt, wie die Forschungsevidenz selektiert oder hergestellt wurde				x			1
beschreibt die Qualität der verwendeten Forschungsevidenz				x			1
gibt Evidenz, dass die PDA die Übereinstimmung zwischen den Präferenzen des informierten Pat. und der gewählten Option verbessert				x			1
gibt Evidenz, dass PDA den Patienten ihr Wissen über Eigenschaften der Optionen verbessert				x			1

Tabellen zur Auflistung Qualitätskriterien

	DISCERN	EQIP 36	HON	IPDASI 4.0	PEMAT	LJIDA	SAM
Beschreibung, was im Test gemessen werden soll				x			I
Wenn der Test das Problem entdeckt, beschreibt die PDA die nächsten typischen Schritte, die gemacht werden				x			I
beschreibt die nächsten Schritte, wenn der Zustand oder die Problematik nicht entdeckt wird				x			I
hat Info über Konsequenz der Entdeckung des Zustands oder Kh, die nie Probleme verursacht hätte, wenn das Screening nicht durchgeführt worden wäre				x			I
Information über Chancen ein richtig-positives Testergebnis zu haben				x			I
Information über Chancen ein richtig-negatives Testergebnis zu haben				x			I
Information über Chancen ein falsch-positives Testergebnis zu haben				x			I
Information über Chancen ein falsch-positives Testergebnis zu haben				x			I
beschreibt Wahrscheinlichkeit, dass die Krankheit mit oder ohne Benutzung des Tests entdeckt wurde				x			I
medizinische Begriffe werden nur verwendet, um Leser mit den Begriffen vertraut zu machen;ansonsten klar definiert					x		I
Zahlen sind klar und einfach zu verstehen					x		I
keine Erwartungen an Benutzer/Benutzer Berechnungen durchzuführen					x		I
GI unterteilt in kürzere Abschnitte					x		I
verwendet visuelle Hinweise (Pfeile, Kästen etc.)					x		I
Text auf Bildschirm einfach zu lesen					x		I
verwendet visuelle Hilfsmittel, wenn man dadurch Inhalt leichter versteht					x		I
Material erlaubt Benutzerin/Benutzer Wörter klar zu hören (nicht zu schnell)					x		I
bietet Werkzeug (z.B. Checklisten) wenn es Benutzerinnen/Benutzer helfen kann Maßnahme zu machen					x		I
bietet einfache Anweisungen oder Beispiel, wie man Berechnungen durchführt					x		I
erklärt wie Diagramme, Schaubilder, Tabellen verwendet werden können, um Maßnahmen zu treffen					x		I
verwendet visuelle Hilfsmittel, wenn sie das Befolgen der Anweisungen erleichtern können					x		I
Ist der Detaillierungsgrad dem Wissensstand der Benutzerin/Benutzer angemessen?						x	I
Kannst du immer den aktuellen Standort auf der Website erkennen?						x	I
Farbschema angemessen und ansprechend						x	I
gleiches Seitenlayout auf der gesamten Seite						x	I
Haben die Navigationslinks eine durchgängig gleichbleibende Funktion?						x	I
Struktur durchgängig gleich						x	I
Website effektive Suchfunktion?						x	I
effektive Browsing Möglichkeiten?						x	I
Minimiert das Design den kognitiven Aufwand bei der Nutzung der Seite						x	I
Unterstützt die Seite normale Browser Navigationswerkzeuge						x	I
Kann die Seite ohne Plug-ins von Drittanbietern verwendet werden						x	I
initiale Definition, welche Themen abgedeckt sind		x					I

Tabelle 9 Aufzählung Qualitätskriterien

3.7 Überprüfung der Instrumente auf die 15 Qualitätskriterien

Wie schon in Kapitel 1.3 beschrieben wurden von der Österreichischen Plattform Gesundheitskompetenz (ÖPGK) in Kooperation mit dem Bundesministerium für Gesundheit und Frauen (BMGF) eine Richtlinie für das Erstellen von guten Gesundheitsinformationen (8) erstellt. Diese soll eine Erleichterung bei der Erstellung von zuverlässiger GI schaffen. Nun wurden im nächsten Schritt geprüft, wie viele der Qualitätskriterien der Richtlinie (8) bei den einzelnen Instrumenten zutreffen. Dabei wurden wieder die Instrumente untersucht, welche in Kapitel 3.4 schon genauer durchleuchtet wurden (DISCERN (16), EQIP36 (26), HON (19), IPDASi4.0 (33), PEMAT (34), LIDA (22), SAM (27)). Falls ein Kriterium nur teilweise bei den einzelnen Instrumenten behandelt wird, wurde es im Rahmen der nachfolgenden Analyse zur Hälfte gewertet.

Die Analyse ergab, dass kein Instrument alle Qualitätskriterien beinhaltet. Das Instrument mit den meisten enthaltenen Kriterien ist das IPDASi4.0 (33). Es enthält zehn der 15 Qualitätskriterien. Das Instrument LIDA (22) beinhaltet sieben der vorhandenen Kriterien. Die Instrumente EQIP36 (26) und HON (19) weisen sechs der 15 Qualitätskriterien auf. Das Instrument DISCERN (16) beinhaltet 5,5 Kriterien. Die Instrumente SAM (27) und PEMAT (34) beinhalten zwei bzw. 1,5 Qualitätskriterien.

Des Weiteren zeigte die Analyse, dass es ein Qualitätskriterium gibt, welches von allen sieben Instrumenten verwendet wird, und ein Qualitätskriterium, welches von keinem einzigen Instrument verwendet wird. Das Qualitätskriterium „Beschreibung der Formate und Inhalte“ wird in allen Instrumenten aufgegriffen. Wobei anzumerken ist, dass in den Instrumenten PEMAT und SAM sogar eine Zusammenfassung erwartet wird. Andererseits wird in keinem der sieben Instrumente die Thematik einer klaren Trennung zwischen Empfehlung und Information aufgegriffen und somit erfüllt kein Instrument das Qualitätskriterium „Bewertungen und Empfehlungen“.

	DISCERN	EQIP 36	HON	IPDASi 4.0	PEM ⁴	SAM
anderer Informationsbedürfnisse						
ercherche	x	x	x	x		
Z						
ng von Ergebnissen (Endpunkten)				x		
ng von Vergleichen				x		
und Risikoangaben	x	x	x	x		
on Alters- und Geschlechterunterschieden	x	x	x	x		
elgruppe				x		x
e Darstellung	(x)	x	x			
mpfehlungen						
stellung von Entscheidungshilfen						
erantwortliche		x	x	x		
essenskonflikten	x					
ormate und Inhalte	x	x	x	x	x	x
nhalte				x		
	5,5	6	6	10	1,	2

Legende
x= 1 Punkt
(x)=0,5
Punkte

Tabelle 10 Überprüfung der 15 Qualitätskriterien in den Instrumenten zur Bewertung der Q

4 Diskussion

Eine Vielzahl von unterschiedlichen Instrumenten zur Bewertung der Qualität von GI wurden durch diese Arbeit ersichtlich. Bei der genaueren Betrachtung dieser stellte sich jedoch heraus, dass die Mehrheit dieser Instrumente entweder nicht vollständig sind oder nur teilweise oder sogar gar nicht validiert sind. Außerdem zeigten sich im Vergleich große Unterschiede zwischen den Instrumenten, wie zum Beispiel bezüglich ihrer Zielgruppe. Dennoch ergab die Auswertung der Anwenderstudien, dass die verschiedenen Instrumente häufig in Anwenderstudien eingesetzt werden.

Identifizierung der Instrumente zur Bewertung der GI

Hinsichtlich der Identifizierung der Instrumente zur Bewertung der GI besteht die Stärke dieser Arbeit vor allem in der Methodik: Zum einen wurde eine breite Literaturrecherche durchgeführt, wodurch dann eine große Bandbreite von Paper resultierte. Zum anderen ist durch das Screening der Ergebnisse im weiteren Schritt auf Abstract- und anschließend auf Volltextebene, welches von zwei Reviewern – im Zweifelsfall von drei – durchgeführt wurde, eine Objektivität und Neutralität gegeben worden. Trotz der breiten Literaturrecherche ergab die Suche jedoch eine recht geringe Anzahl von nur acht Originalarbeiten, welche das Screening auf Abstract- und Volltextscreening bestanden haben und somit nicht ausgeschlossen wurden. Eine mögliche Idee zur Erhöhung der Anzahl der Ergebnisse ist eine noch breiter gefächerte Suche, welche zum Beispiel mehr Datenbanken beinhaltet. Eine Schwäche liegt Arbeit liegt wahrscheinlich darin, dass sowohl Instrumente zur Beurteilung der Lesbarkeit von Texten, sowie auch Webseiten, welche valide GI sammeln, einbezogen wurden. Außerdem wurden die Instrumente für die Beurteilung der Lesbarkeit vor allem in den Reviews oftmals unterschiedlich benannt und somit wurden anfangs viele Instrumente beim Abstract- und Volltextscreening mehrfach genannt. Erst bei genauerer Betrachtung und Beurteilung mancher Instrumente wurde ersichtlich, dass es sich um das gleiche Instrument handelte. Diese Problematik zeigte sich am meisten bei den Reviews, da vor allem in den Reviews in denen eine große Anzahl von Instrumenten gefunden wurde, wie zum Beispiel in dem Review von Breckons et al. (24), oftmals nur Instrumente ohne genauere Beschreibungen aufgezählt wurden.

Beschreibung der Kriterien

Durch die Kriterienliste der Instrumente zur Beurteilung der Qualität von GI wird die Häufigkeit der einzelnen Kriterien in der Arbeit klar ersichtlich. Ebenso werden die Unterschiede der einzelnen Instrumente durch die Kriterienliste genau dargestellt. Weil die

verschiedenen Instrumente teilweise für verschiedene Gesundheitsinformationen entwickelt wurden, verwenden manche Instrumente spezielle Kriterien, welche deshalb nicht auch noch von einem anderen Instrument verwendet werden. Zum Beispiel beschäftigen sich die Kriterien des Instruments LIDA (22) unter anderem auch mit dem Vorhandensein einer effektiven Suchfunktion und Browsing-Möglichkeit, da das Instrument auf Gesundheitswebseiten ausgelegt ist. Da die anderen Instrumente vorwiegend für gedruckte Gesundheitsinformationen erstellt wurde, finden diese Kriterien selbstverständlich keine Anwendung bei den Instrumenten.

Gleichzeitig liegt darin jedoch auch die Schwäche der Kriterienliste: Dadurch dass insgesamt 138 verschiedene Kriterien in den Instrumenten gefunden wurden, ist die Kriterienliste vor allem bezüglich der Kriterien, welche nur in einem Instrument genannt wurden, unübersichtlich. Durch eine unterschiedliche Farbgestaltung der verschiedenen Häufigkeiten der Kriterien wurde die Übersichtlichkeit optimiert. Eine weitere Schwäche der Kriterienliste ist eine hohe Fehleranfälligkeit beim Zusammenlegen von ähnlichen Kriterien verschiedener Instrumente. Dieses Zusammenlegen hätte durch einheitliche, objektive Kriterien vereinheitlicht werden können.

Sichtung der Instrumente bezüglich der 15 Qualitätskriterien

Auch bezüglich der Sichtung der Instrumente in Hinblick auf die 15 Qualitätskriterien (8) gibt es Stärken und Schwächen. Einerseits überzeugt die Arbeit besonders in der genauen Aufschlüsselung und der Übersichtlichkeit der Ergebnisse. Anhand der hier erstellten Tabelle (Tabelle 7) sind die Unterschiede bzw. die verschiedenen Schwerpunkte der Instrumente genau ersichtlich. Allerdings überschneiden sich die Qualitätskriterien des Öfteren und dies erschwert somit eine klare Auswertung, ob die Instrumente die jeweiligen Kriterien erfüllen oder nicht. Deswegen wurden auch halbe Punkte vergeben, um die Schwere der Gewichtung etwas zu reduzieren. Zusätzlich lassen die Ergebnisse vermuten, dass Entscheidungshilfen leichter mehr Qualitätskriterien erfüllen können. Dies ist wahrscheinlich der Fall, weil die Qualitätskriterien der Kategorie „Auswahl und Darstellung der Fakten“ stark von der Thematik der Entscheidungsfähigkeit, die die Leserinnen/ Leser durch die Verwendung der GI bekommen, geprägt ist. Diese Theorie bestätigt sich ebenfalls bei dem Instrument LIDA: Dieses wurde – wie schon erwähnt – für die Erstellung von Gesundheitswebseiten entwickelt und beschäftigt sich somit zwar nicht mit Entscheidungshilfen, aber sollte andererseits die besten Voraussetzungen haben, die Qualitätskriterien dieser Richtlinie zur Erstellung von GI erfüllen zu können. Bei der Auswertung zeigt sich, dass das Instrument LIDA alle Qualitätskriterien außer die der

Kategorie „Auswahl und Darstellung der Fakten“ erreicht und somit lediglich auf sieben Punkte kommt. Des Weiteren ist das Kriterium „Vorgehen bei der Erstellung von Entscheidungshilfen“ generell nur für Patientinnen/ Patienten-Entscheidungshilfs-Instrumente. Somit kann nur das Instrument IPDASi4.0 (33) - als einziges Instrument für die Bewertung von Entscheidungshilfen – dieses Qualitätskriterium vollständig erfüllen.

Grundsätzlich ist die Problematik bezüglich guter GI und guten Instrumenten zur Bewertung der Qualität von GI dieselbe: Sowohl bei den GI als auch bei den Instrumenten gibt es eine immer größer werdende Zahl und somit ist es für die Benutzerin/ Benutzer schwierig zu erkennen, welche GI bzw. welches Instrument valide ist. Bezüglich der GI wurde im MEDIB-Projekt (6) – wie schon in der Einleitung erwähnt - festgestellt, dass die bestehenden GI in den Hausarztpraxen in der Steiermark nicht den Qualitätskriterien entsprechen, obwohl schon Qualitätskriterien für die Erstellung von guter GI entwickelt wurden. In der Diplomarbeit wird die Vielzahl von Instrumenten ersichtlich und zudem zeigte sich, dass viele Instrumente schon bei einer oberflächlichen Beurteilung nicht den Ausschlusskriterien standhalten können und nur eine geringe Anzahl diesen Ausschlusskriterien überhaupt standhalten kann. Ein weiteres Problem, mit dem Benutzerinnen/ Benutzer konfrontiert sind, ist die eingeschränkte Alltagstauglichkeit und Handhabung der meisten Instrumente. Grundsätzlich sollte ein Instrument möglichst wenig Zeit benötigen, um die Qualität einer GI zu beurteilen. Instrumente wie zum Beispiel das Instrument IPDASi (33) sind mit 44 Items sehr umfangreich und sind daher fraglich bezüglich ihrer Alltagstauglichkeit. Sehr praktisch ist hierbei das Browser Plugin von den Entwicklern des Instruments HON (19), mit dem bei der Internetsuche von medizinischen Themen Websites gekennzeichnet werden, welche den HON Richtlinien entsprechen. Dies erleichtert schnell die Identifizierung von guten GI und ist auch für ältere Personengruppen gut geeignet, welche zwar im Internet nach GI suchen, aber für die Anwendung der anderen Instrumente zu schwierig ist. Trotzdem zeigt sich die Dringlichkeit entweder eines validen Instruments zur Beurteilung der Qualität von GI oder einer Plattform, welche für die Überprüfung von GI zuständig ist, damit Benutzerinnen/ Benutzer einfach und verlässlich gute GI finden können.

Eine übergreifende Limitation in dieser Arbeit ist vor allem der fehlende Einbezug der Patientinnen/ Patienten bezüglich deren Standpunkt gegenüber den Instrumenten zur Bewertung der Qualität von GI: Einerseits fehlt die Erfahrung von Patientinnen/ Patienten bezüglich der Alltagstauglichkeit dieser Instrumente. Andererseits stellt sich die Frage, wie groß die Nachfrage für ein Instrument von Seiten der Patientinnen/ Patienten ist und ob diese nicht eher von Seiten der Allgemeinmedizinerinnen/ Allgemeinmedizinern besteht und im

weiteren Sinne die Vermittlung von guter GI die Aufgabe von Allgemeinmedizinerinnen/ Allgemeinmedizinern ist.

Ein weiterer wichtiger Punkt ist, dass die derzeitige Problematik immer noch darin liegt, dass die meisten GI in Hausarztpraxen (6) nicht valide sind. Somit muss weiterhin – zum Beispiel durch Initiativen und Projekte – die Erstellung von validen GI unterstützt werden, damit die Zahl von validen GI steigt. Außerdem braucht es ein validiertes Instrument, das unterschiedlichen Personengruppen eine Einschätzung der Qualität von GI erlaubt.

Eine solche Checkliste (MAPPinfo. Mapping quality of health information) (15), die ohne besondere Schulung oder andere Hilfsmittel eingesetzt werden kann, wurde bereits entwickelt und wird momentan validiert.

Auch in Zukunft werden sich GI immer weiter vermehren und folglich wird sich auch die Situation für Patientinnen/ Patienten weiter erschweren. Ebenfalls werden sich immer mehr Patientinnen/ Patienten bereits vor dem Arztbesuch anderweitig informieren, wodurch sie auf eine Flut von GI treffen. Daher ist es auf der einen Seite wichtig die Gesundheitskompetenz der Patientinnen/ Patienten, zum Beispiel durch Schulungen und Gesundheitsprogramme, zu stärken. Auf der anderen Seite benötigt es eine unabhängige Stelle, welche evidenzbasierte GI sammelt und bewertet. Dadurch hätten auch Ärztinnen/ Ärzte oder auch andere Personen in Gesundheitsberufen eine verlässliche Möglichkeit ihren Patientinnen/ Patienten valide GI mitzugeben, ohne dass dies viel Zeit in Anspruch nehmen würde. Somit würde allen beteiligten Parteien geholfen werden.

Literaturverzeichnis

1. Marstedt G. Das Internet: Auch Ihr Ratgeber für Gesundheitsfragen? [Stand: 11.03.2020]. Verfügbar unter: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/VV_Studie_Das-Internet-auch-Ihr-Ratgeber_Befragung.pdf.
2. Hambrock U. Die Suche nach Gesundheitsinformationen; 2018.
3. Mueller J, Jay C, Harper S, Todd C. The Role of Web-Based Health Information in Help-Seeking Behavior Prior to a Diagnosis of Lung Cancer: A Mixed-Methods Study. *J Med Internet Res* 2017; 19(6):e189. doi: 10.2196/jmir.6336.
4. Tan SS-L, Goonawardene N. Internet Health Information Seeking and the Patient-Physician Relationship: A Systematic Review. *J Med Internet Res* 2017; 19(1):e9. doi: 10.2196/jmir.5729.
5. Kerschner B, Wipplinger J, Klerings I, Gartlehner G. Wie evidenzbasiert berichten Print- und Online-Medien in Österreich? Eine quantitative Analyse. *Z Evid Fortbild Qual Gesundhwes* 2015; 109(4-5):341–9. doi: 10.1016/j.zefq.2015.05.014.
6. IAMEV. Medizinische Informationsbroschüren in der Hausarztpraxis (MEDIB) [Stand: 02.04.2020]. Verfügbar unter: https://allgemeinmedizin.medunigraz.at/fileadmin/institute-oes/allgemeinmedizin/Publikationen/Berichte/2017/IAMEV_MEDIB_final.pdf.
7. European Commission; Directorate-General for Health and Consumers. White Paper: Together for Health: a Strategic Approach for the EU 2008-2013. 1. Aufl. Saarbrücken: Dictus Publishing; 2011.
8. Arbeitsgruppe Gute Gesundheitsinformation (AG GGI) der Österreichischen Plattform Gesundheitskompetenz (ÖPGK). Überblick über die 15 Qualitätskriterien für zielgruppenorientierte, evidenzbasierte Broschüren, Videos, Websites und Apps; 2020.
9. Sørensen K, van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z et al. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health* 2012; 12:80. doi: 10.1186/1471-2458-12-80.

10. Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G et al. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). *Eur J Public Health* 2015; 25(6):1053–8. doi: 10.1093/eurpub/ckv043.
11. Zielsteuerungsvertrag 2017 bis 2021; 2020 [Stand: 11.04.2020]. Verfügbar unter: [https://www.sozialministerium.at/Themen/Gesundheit/Gesundheitssystem/Gesundheitsreform-\(Zielsteuerung-Gesundheit\)/Zielsteuerungsvertrag-2017-bis-2021.html](https://www.sozialministerium.at/Themen/Gesundheit/Gesundheitssystem/Gesundheitsreform-(Zielsteuerung-Gesundheit)/Zielsteuerungsvertrag-2017-bis-2021.html).
12. Bundesministerium für Gesundheit und Frauen (BMGF). 10 Ziele - Gesundheitsziele Österreich; 2020 [Stand: 11.04.2020]. Verfügbar unter: <https://gesundheitsziele-oesterreich.at/10-ziele/>.
13. Arbeitsgruppe GPPI. Gute Praxis Gesundheitsinformation. *Z Evid Fortbild Qual Gesundheitswes* 2016; 110-111:85–92. Verfügbar unter: <http://www.sciencedirect.com/science/article/pii/S1865921715300015>.
14. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* 2006; 333(7565):417. doi: 10.1136/bmj.38926.629329.AE.
15. Kasper J, Lühnen J, Hinneburg J, Siebenhofer A, Posch N, Berger-Höger B et al. MAPPinfo, mapping quality of health information: study protocol for a validation study of an assessment instrument. *BMJ Open* 2020; 10(11):e040572. doi: 10.1136/bmjopen-2020-040572.
16. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999; 53(2):105–11. doi: 10.1136/jech.53.2.105.
17. Daraz L, MacDermid JC, Wilkins S, Shaw L. Tools to Evaluate the Quality of Web Health Information: A Structured Review of Content and Usability. *The International Journal of Technology, Knowledge, and Society* 2009; 5(3):127–42. doi: 10.18848/1832-3669/cgp/v05i03/55997.
18. Kincaid JP, Fishburne RP, JR., Rogers RL, and Chissom BS. Derivation Of New Readability Formulas (Automated Readability Index, Fog Count And Flesch Reading Ease Formula) For Navy Enlisted Personnel. Millington (Tenn.); 1975.

19. Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health Websites. *Comput Biol Med* 1998; 28(5):603–10. doi: 10.1016/s0010-4825(98)00037-7.
20. FLESCH R. A new readability yardstick. *J Appl Psychol* 1948; 32(3):221–33. doi: 10.1037/h0057532.
21. JAMA benchmarks/standards [Stand: 24.08.2021]. Verfügbar unter: <https://jamanetwork.com/>.
22. Minervation Limited. Welcome to the LIDA Blog [Stand: 23.08.2021]. Verfügbar unter: <http://www.minervation.com/welcome-to-the-lida-blog/>.
23. G. Harry Mc Laughlin. SMOG Grading — a New Readability Formula; 1969 8.
24. Breckons M, Jones R, Morris J, Richardson J. What do evaluation instruments tell us about the quality of complementary medicine information on the internet? *J Med Internet Res* 2008; 10(1):e3. doi: 10.2196/jmir.961.
25. Luk A, Aslani P. Tools used to evaluate written medicine and health information: document and user perspectives. *Health Educ Behav* 2011; 38(4):389–403. doi: 10.1177/1090198110379576.
26. Charvet-Berard AI, Chopard P, Perneger TV. Measuring quality of patient information documents with an expanded EQIP scale. *Patient Educ Couns* 2008; 70(3):407–11. doi: 10.1016/j.pec.2007.11.018.
27. Doak, C. C., Doak, L. G. & Root, J. H. Teaching Patients with Low Literacy Skills. 2. Aufl. Philadelphia: J. B. Lippincott Company; 1996.
28. Khazaal Y, Chatton A, Cochand S, Coquard O, Fernandez S, Khan R et al. Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. *Patient Educ Couns* 2009; 77(1):33–7. doi: 10.1016/j.pec.2009.02.016.
29. National Institutes of Health. MedlinePlus: Trusted Health Information for you; 2021 [Stand: 24.03.2021]. Verfügbar unter: <https://medlineplus.gov/>.
30. AMA [Stand: 23.08.2021]. Verfügbar unter: <https://www.ama-assn.org/>.
31. Gunning R. The technique of clear writing. Rev. ed. New York: McGraw-Hill; 1968.

32. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor--Let the reader and viewer beware. *JAMA* 1997; 277(15):1244–5.
33. Joseph-Williams N, Newcombe R, Politi M, Durand M-A, Sivell S, Stacey D et al. Toward Minimum Standards for Certifying Patient Decision Aids: A Modified Delphi Consensus Process. *Med Decis Making* 2014; 34(6):699–710. doi: 10.1177/0272989X13501721.
34. Shoemaker SJ, Wolf MS, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ Couns* 2014; 96(3):395–403. doi: 10.1016/j.pec.2014.05.027.
35. Moulton B, Franck LS, Brady H. Ensuring quality information for patients: development and preliminary validation of a new instrument to improve the quality of written health care information. *Health Expect* 2004; 7(2):165–75. doi: 10.1111/j.1369-7625.2004.00273.x.
36. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* 2006; 333(7565):417. doi: 10.1136/bmj.38926.629329.AE.
37. Elwyn G, O'Connor AM, Bennett C, Newcombe RG, Politi M, Durand M-A et al. Assessing the quality of decision support technologies using the International Patient Decision Aid Standards instrument (IPDASi). *PLoS ONE* 2009; 4(3):e4705. doi: 10.1371/journal.pone.0004705.
38. Amstad T. *Wie verständlich sind unsere Zeitungen?* [Diss. phil. I Zürich]. S.l.: s.n; 1978.
39. Contreras A, García-Alonso R, Echenique M, Daye-Contreras F. The SOL formulas for converting SMOG readability scores between health education materials written in Spanish, English, and French. *J Health Commun* 1999; 4(1):21–9. doi: 10.1080/108107399127066.

Anhang

Anhang1: Suchstrategie

A) Suche in OVID (nur Medline)

Datum: 10.07.2019

Durchführung: kjj/np

Such-Set:

Datenbankstatus: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) 1946 to July 08, 2019

Recherchezeitraum: bis dato

Filter: keine

Anmerkungen: gespeichert als „190710_Mini-Checkliste_GI“

MeSH-Terms und Textworte

Bestimmung der relevanten Begriffe/MeSH-Terms erfolgte im Wesentlichen in 2 Schritten:

1. „patient information“ in pubmed > 400 Treffer > Durchsicht dieser nach MeSh terms
2. Ausgehend von 4 bereits bekannten relevanten Publikationen (=Test-Set) > related articles in pubmed > Durchsicht der Abstracts nach MeSH terms und Textanalyse der Abstracts (Durchsicht nach Textworten und deren Kombinationen)

adj5

Die Verwendung von "adj5" statt "and" im Block "Qualitätsbewertung/Kriterien" war erforderlich, weil ansonsten die Trefferzahl weit über 60.000 gelegen wäre. Der Wortabstand von "5" ist willkürlich gewählt.

Arbeitshypothese: jeweils 2 Begriffe aus den nachfolgenden 3 Bereichen müssen innerhalb von 5 Worten zu finden sein:

- (instrument\$ or criteria or item\$ or standard\$ or tool\$ or manual\$ or checklist\$)
- (assess\$ or measur\$ or rating or rate or apprais\$ or evaluat\$ or validat\$ or reliab\$)
- quality

Auch die MeSH-Terms im Block "Qualitätsbewertung/Kriterien" müssen einen Bezug zu "Qualität" aufweisen.

	Searches	Results
1	Health Education/	59077
2	exp Consumer Health Information/	8040
3	Patient Education as Topic/	82302
4	exp Teach-Back Communication/	27
5	Patient Education Handout.pt.	4944
6	Decision Making/	89368
7	Decision Support Techniques/	18966
8	Pamphlets/	3758
9	exp Patient Participation/	24080
10	or/1-9	270671
11	(patient\$ or consumer\$).ti,ab.	6336657
12	information.ti,ab.	1126164
13	(decision\$ and (aid\$ or support\$ or making)).ti,ab.	177732
14	education material\$.ti,ab.	1757
15	leaflet\$.ti,ab.	21019
16	11 and (12 or 13 or 14 or 15)	422624
17	health Information.ti,ab.	20359
18	10 or 16 or 17 [Block: Patienteninformation/Entscheidungshilfe/Gesundheitsinformation]	663518
19	exp Checklist/	5620
20	exp Manuals as Topic/	3701
21	Quality Assurance, Health Care/	55011
22	(19 or 20) and (21 or quality.ti,ab.)	1450
23	((instrument\$ or criteria or item\$ or standard\$ or tool\$ or manual\$ or checklist\$) adj5 quality).ti,ab.	49252
24	((assess\$ or measur\$ or rating or rate or apprais\$ or evaluat\$ or validat\$ or reliab\$) adj5 quality).ti,ab.	179162

25	((instrument\$ or criteria or item\$ or standard\$ or tool\$ or manual\$ or checklist\$) adj5 (assess\$ or measur\$ or rating or rate or apprais\$ or evaluat\$ or validat\$ or reliab\$)).ti,ab.	391566
26	23 and 24 and 25 [Block: Qualitätsbewertung/Kriterien]	14409
27	22 or 26	15709
28	18 and 27	1876
29	("10396471" or "18242935" or "19259269" or "15117391"an. [Test-Set]	4
30	28 and 29	4

C) Suche in PsycINFO

Datum: 18.07.2019

Durchführung: kw/kjj/np

Such-Set:

Datenbankstatus: PsycINFO 1806 to July Week 2 2019

Recherchezeitraum: bis dato

Filter: keine

Anmerkungen: gespeichert als „190718_Mini-Checkliste_GI“ (KW-Account)

MeSH-Terms und Textworte

Analoges Vorgehen zu Medline

adj5

Analoges Vorgehen zu Medline

	Suchen	Ergebnisse
1	health information.sh.	2024
2	health literacy.sh.	2660
3	1 or 2	4483
4	(patient\$ or consumer\$).ti,ab.	711517
5	information.ti,ab.	415157

6	(decision\$ and (aid\$ or support\$ or making)).ti,ab.	112626
7	education material\$.ti,ab.	599
8	leaflet\$.ti,ab.	890
9	4 and (5 or 6 or 7 or 8)	83725
10	health Information.ti,ab.	5556
11	3 or (9 or 10)	89279
12	Test Construction.sh.	42212
13	Test validity.sh.	74111
14	test reliability.sh.	52320
15	12 or 13 or 14	109507
16	Quality of Services.sh.	6079
17	Quality Control.sh.	1455
18	16 or 17 or quality.ti,ab.	222730
19	15 and 18	8846
20	((instrument\$ or criteria or item\$ or standard\$ or tool\$ or manual\$ or checklist\$) adj5 quality).ti,ab.	9918
21	((assess\$ or measur\$ or rating or rate or apprais\$ or evaluat\$ or validat\$ or reliab\$) adj5 quality).ti,ab.	37272
22	((instrument\$ or criteria or item\$ or standard\$ or tool\$ or manual\$ or checklist\$) adj5 (assess\$ or measur\$ or rating or rate or apprais\$ or evaluat\$ or validat\$ or reliab\$)).ti,ab.	153733
23	20 and 21 and 22	3477
24	19 or 23	11430
25	11 and 24	731