

# Stehen die Bewertungen von Individuellen Gesundheitsleistungen (IGeL) des IGeL-Monitors im Einklang mit Leitlinien?

## Are the assessments of individual out-of-pocket health services of the IGeL-Monitor in line with clinical guidelines?



Autorinnen/Autoren

Monika Becker, Ute Hansen, Michaela Eikermann

### Institute

Bereich Evidenzbasierte Medizin, Medizinischer Dienst Bund, Essen, Germany

### Schlüsselwörter

IGeL, Individuelle Gesundheitsleistungen, IGeL-Monitor, Leitlinien, Evidenzbasierte Medizin

### Key words

IGeL, individual out-of-pocket health services, IGeL-Monitor, guidelines, evidence-based medicine

### Bibliografie

Gesundheitswesen 2023; 85: 1192–1199

DOI 10.1055/a-2158-8869

ISSN 0941-3790


© 2023. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

### Korrespondenzadresse

Monika Becker  
Medizinischer Dienst Bund,  
Bereich Evidenzbasierte Medizin,  
Theodor-Althoff-Str. 47  
45133 Essen  
Germany  
[monika.becker@md-bund.de](mailto:monika.becker@md-bund.de)

 **Zusätzliches Material** finden Sie unter <https://doi.org/10.1055/a-2011-5362>

### ZUSAMMENFASSUNG

**Ziel der Studie** Der IGeL-Monitor des Medizinischen Dienstes Bund untersucht Nutzen und Schaden Individueller Gesund-

heitsleistungen (IGeL). Das Ziel der Analyse war eine systematische Gegenüberstellung von Bewertungen aktueller IGeL-Themen und den Empfehlungen aus evidenzbasierten Leitlinien.

**Methodik** Zur Identifikation evidenzbasierter Leitlinien wurden Recherchen in Leitliniendatenbanken (AWMF, Guidelines International Network und Trip Database) sowie auf Internetseiten von Leitlinienorganisationen durchgeführt (Februar/März 2022). Es wurden Leitlinien eingeschlossen, die nicht älter als 5 Jahre waren. Die methodische Qualität der Leitlinien wurde mit dem AGREE II-Instrument bewertet. Die Empfehlungen wurden inhaltlich sowie hinsichtlich der Empfehlungsstärke mit den IGeL-Bewertungen abgeglichen.

**Ergebnisse** Es wurden 41 LL zu 24 aktuellen IGeL-Themen identifiziert. 19 (79 %) Bewertungen stimmten (nahezu) mit den Leitlinienempfehlungen überein. Zu fünf IGeL-Themen war kein Abgleich möglich, da zum Beispiel die Empfehlungen spezifischer waren. Zehn der 13 IGeL, die (tendenziell) negativ bewertet wurden, wurden auch in den Leitlinien nicht empfohlen.

**Schlussfolgerung** In der Gesamtschau stimmen die Aussagen aus den IGeL-Bewertungen mit den Empfehlungen aktueller LL überein. Hiernach scheinen Leitliniengruppen die Evidenz ähnlich einzuschätzen wie das Team des IGeL-Monitors. Insbesondere zu (tendenziell) negativ bewerteten IGeL, die auch in Leitlinien nicht empfohlen werden, sollten Versicherte ehrlich über die Evidenz aufgeklärt werden.

### ABSTRACT

**Objectives** The IGeL-Monitor of the Federal Medical Advisory Service in Germany evaluates the benefits and harms of individual out-of-pocket health services (in German: Individuelle Gesundheitsleistungen / IGeL). The aim of the analysis was to systematically compare IGeL-assessments with the recommendations from evidence-based guidelines.

**Method** To identify guidelines, we conducted searches in guidelines databases (AWMF, Guidelines International Network and Trip database) and the websites of guideline organisations (February/March 2022). We included guidelines that were not older than 5 years. The methodological quality of the guidelines was assessed using the AGREE II instrument. We compared

the recommendations with the IGeL-assessments in terms of content and grade of recommendation.

**Results** We identified 41 guidelines covering 24 IGeL-assessments. 19 (79%) assessments (nearly) were in agreement with the guideline recommendations. No comparison was possible for 5 IGeL-assessments, because, for example, the recommendations were more specific. Ten of the 13 IGeL that were rated (tendentially) negatively were also not recommended in the guidelines.

**Conclusion** Overall, the IGeL-assessments were consistent with the recommendations of current guidelines. Accordingly, guideline groups seem to assess the evidence similarly to the IGeL-Monitor team. Insured persons should be informed honestly about the evidence, particularly for the (tendentially) negatively evaluated IGeL that are not recommended even in guidelines.

## Einleitung

Der IGeL-Monitor des Medizinischen Dienstes Bund untersucht seit 2012 Nutzen und Schaden Individueller Gesundheitsleistungen (IGeL). Unter IGeL sind Leistungen zu verstehen, die nicht zum Leistungskatalog der gesetzlichen Krankenkassen gehören, da beispielsweise keine ausreichenden Belege für ihren Nutzen vorliegen [1]. Versicherte müssen für diese Leistungen in der Regel selbst zahlen. Gesetzlich Krankenversicherte in Deutschland geben schätzungsweise etwa 1 Milliarde Euro pro Jahr für IGeL aus [2]. Ziel des IGeL-Monitors ist es, Versicherte bei ihrer Entscheidung für oder gegen eine IGeL zu unterstützen. Bisher wurden 53 Bewertungen erstellt (Stand November 2022), darunter die am häufigsten von Ärzt\*innen angebotenen IGeL, wie Augeninnendruckmessung zur Glaukom-Früherkennung sowie Ultraschall der Eierstöcke und der Brust zur Krebsfrüherkennung [3].

Die Bewertungen des IGeL-Monitors basieren auf einer systematischen Aufbereitung von systematischen Übersichtsarbeiten und / oder Primärstudien. Die Ergebnissicherheit und Größe der studienübergreifenden Ergebnisse zu patientenrelevanten Nutzen- und Schadenendpunkten werden jeweils einer von drei Kategorien zugeordnet: kein Hinweis / Hinweis / Beleg. Ein Beleg für einen Nutzen oder Schaden kann abgeleitet werden bei Vorliegen einer hohen Ergebnissicherheit und konsistenten Ergebnissen aus qualitativ hochwertigen Studien. Bei weniger sicheren und konsistenten Ergebnissen wird ein Hinweis abgeleitet. Für die Gesamtbewertung werden Nutzen und Schaden gegeneinander abgewogen und die einzelnen Leistungen in fünf Abstufungen wie folgt bewertet:

- Positiv: Belege für Nutzen und keine Hinweise auf Schaden
- Tendenziell positiv: Hinweise auf Nutzen und keine Hinweise auf Schaden oder Belege für Nutzen und Hinweise auf Schaden
- Unklar: keine Hinweise auf Nutzen und Schaden oder Hinweise auf Nutzen und Schaden oder Belege für Nutzen und Schaden
- Tendenziell negativ: Hinweise auf Schaden und keine Hinweise auf Nutzen oder Belege für Schaden und Hinweise auf Nutzen
- Negativ: Belege für Schaden und keine Hinweise auf Nutzen.

Laut IGeL-Report 2020 werden 80% der IGeL von Ärzt\*innen angeboten, nur 19% werden von Versicherten selbst nachgefragt [3]. Etwa jede zweite der angebotenen IGeL wird von den Versicherten auch in Anspruch genommen [3].

Ergänzend wird im Rahmen der Bewertungen des IGeL-Monitors regelhaft fokussiert nach medizinischen Leitlinien (LL) gesucht, um

die Bewertungen auch in den Kontext aktueller LL-Empfehlungen einordnen zu können. LL bilden den gegenwärtigen wissenschaftlichen Erkenntnisstand ab und sollen die Entscheidungsfindung von Ärzt\*innen und Angehörigen von anderen Gesundheitsberufen sowie Patient\*innen und Bürger\*innen hinsichtlich spezifischer Gesundheitsprobleme unterstützen [4]. Evidenzbasierte LL basieren im Unterschied zu nicht-evidenzbasierten LL, wie zum Beispiel Handlungsempfehlungen von Expertengruppen, auf einer systematischen Recherche und Aufbereitung der Evidenz sowie einer Bewertung von Nutzen und Schaden alternativer Vorgehensweisen [5]. In Deutschland werden LL primär von den wissenschaftlich medizinischen Fachgesellschaften unter dem Dach der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) erstellt. Die höchsten methodischen Anforderungen erfüllen sogenannte S3-LL, welche sich durch eine systematische Aufbereitung der Evidenz, eine formale Konsensfindung und ein repräsentativ besetztes Gremium auszeichnen [4]. Evidenzbasierten LL wird bei der Sicherstellung einer qualitativ hochwertigen Gesundheitsversorgung eine wichtige Rolle zugeschrieben [6, 7]. Verschiedene Evaluationen zur Umsetzung von LL geben in diesem Zusammenhang auch Hinweise auf den Nutzen einer leitliniengerechten Behandlung bezogen auf patientenrelevante Endpunkte [8, 9].

Die Bewertungen des IGeL-Monitors werden vielfach dahingehend kritisiert, dass medizinisch sinnvolle Leistungen diskreditiert würden [10, 11]. So würden auch viele der IGeL in LL empfohlen [11]. Bislang fehlt jedoch eine systematische Gegenüberstellung von aktuellen Empfehlungen aus evidenzbasierten LL und den Bewertungen des IGeL-Monitors.

Das Ziel der Analyse war ein Vergleich der Bewertungen von aktuellen IGeL-Themen mit Empfehlungen aus evidenzbasierten LL hinsichtlich Übereinstimmungen und Abweichungen.

## Methodik

Zur Identifikation relevanter LL wurden Recherchen in den folgenden LL-Datenbanken durchgeführt (Recherchedatum: Februar / März 2022): Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Guidelines International Network (GIN) und Trip Database. Darüber hinaus wurden die Internetseiten der fachübergreifenden LL-Anbieter National Institute for Health and Care Excellence (NICE), Scottish Intercollegiate Guidelines Network (SIGN) und U.S. Preventive Services Task Force (USPSTF) gescreent. Zusätzlich wurden die Referenzlisten relevanter LL geprüft.

In die Analyse wurden alle IGeL-Themen einbezogen, die eine Bewertung beinhalten. Archivierte und in Aktualisierung befindliche Bewertungen wurden nicht berücksichtigt. Für den Vergleich der Bewertungen mit LL-Empfehlungen wurden LL eingeschlossen, die folgende Kriterien erfüllten:

- LL enthält Empfehlung(en) zu mindestens einer der bewerteten IGeL;
- Empfehlungen beziehen sich auf die Versorgung in einem OECD-Mitgliedsstaat;
- Es handelt sich um eine evidenzbasierte LL (systematische Literaturrecherche, Empfehlungen sind mit einem Evidenzlevel und / oder Empfehlungsgrad gekennzeichnet);
- nachvollziehbares Klassifizierungsschemata der Evidenz- und / oder Empfehlungseinstufung;
- nicht älter als 5 Jahre;
- als gültig gekennzeichnet / Überarbeitungsdatum nicht überschritten;
- Publikationssprache: Deutsch oder Englisch;
- Vollpublikation ist frei verfügbar.

In Bezug auf deutsche LL wurden auch nicht-evidenzbasierte LL (informell bzw. formal konsentrierte S1- bzw. S2k-LL) und / oder ältere (> 5 Jahre) LL eingeschlossen. Diese wurden für den deutschen Versorgungskontext gleichwohl als relevant angesehen, insbesondere, da in der Diskussion um IGeL-Bewertungen auch auf solche LL regelhaft Bezug genommen wird.

Das Titel- und Abstractscreening erfolgte durch eine Person. Potentiell relevante Volltexte wurden durch zwei Personen unabhängig hinsichtlich der Einschlusskriterien geprüft. Uneinigkeiten in Bezug auf den Ein- oder Ausschluss eines Volltextes wurden durch Diskussion gelöst.

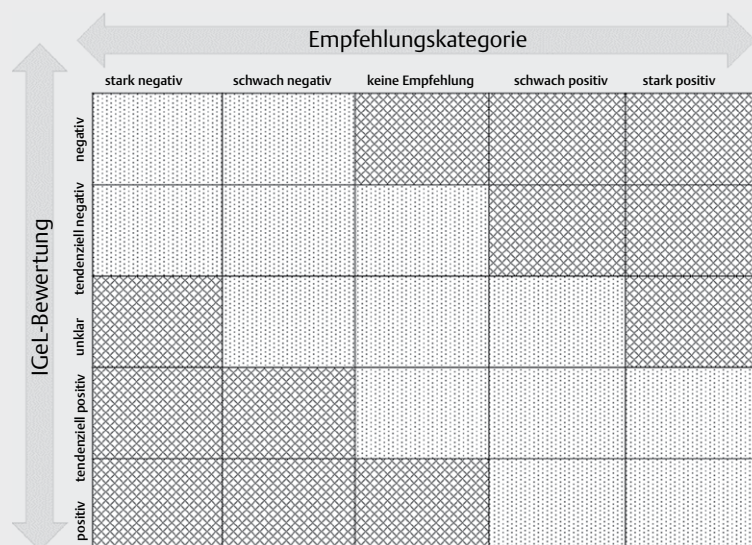
Die Empfehlungen der eingeschlossenen LL wurden mit Evidenzlevel und Empfehlungsgrad in standardisierte Tabellen extrahiert. Da

sich in LL häufig unterschiedliche Systeme der Evidenz- und Empfehlungsgraduierung finden, wurden die in den LL verwendeten Evidenz- und Empfehlungsgrade Kategorien zugeordnet. Für die Evidenzlevel wurden dabei die Kategorien „hoch“, „mittel“ oder „niedrig“ unterschieden. Die Definitionen der Evidenzkategorien finden sich im *Anhang A* (online verfügbar). Bei den Empfehlungsgraden wurden die Empfehlungen mit dem höchsten in der LL verwendeten Empfehlungsgrad der Kategorie „stark“ zugeordnet, alle niedrigeren Empfehlungsgrade der Kategorie „schwach“. Wenn explizit keine Empfehlung für oder gegen eine Intervention ausgesprochen wurde, wurden diese in die Kategorie „keine“ eingeordnet.

Die Empfehlungen wurden inhaltlich sowie hinsichtlich der Empfehlungsstärke mit den IGeL-Bewertungen (positiv/negativ, tendenziell positiv/negativ, unklar) abgeglichen. Unter Einbeziehung der Empfehlungsrichtung konnten sich die Empfehlungen zwischen null und vier Stufen von den IGeL-Bewertungen unterscheiden. Ein Unterschied von null Stufen oder einer Stufe wurde als „(nahezu) Übereinstimmung“ zwischen Empfehlung und IGeL-Bewertung gewertet, ein Unterschied von zwei Stufen oder mehr als „Abweichung“ (siehe ► **Abb. 1**). Sofern einer Empfehlung kein eindeutiger Empfehlungsgrad zugeordnet werden konnte, wurden die Evidenzlevel für den Abgleich verwendet. Dabei wurde ein mittlerer oder niedriger Evidenzlevel wie ein schwacher Empfehlungsgrad gehandhabt.

Die Extraktionen und Abgleiche mit den IGeL-Bewertungen wurden durch eine Person durchgeführt und durch eine zweite Person qualitativ gesichert.

Zur Bewertung der methodischen Qualität der eingeschlossenen LL wurde das Appraisal of Guidelines for Research & Evaluation (AGREE) II-Instrument eingesetzt [12]. Die Anwendung des Instrumentes beschränkte sich auf die Domänen 3 (Genauigkeit der LL-Entwicklung) und Domäne 6 (redaktionelle Unabhängigkeit). Beide Domänen können als besonders relevant für die Vertrauenswür-



► **Abb. 1** Empfehlungskategorien und IGeL-Bewertungen; Punkte = „(nahezu) Übereinstimmung“ bei 0 bis 1 Stufen Unterschied; Linien = „Abweichung“ bei 2 oder mehr Stufen Unterschied.

digkeit der Empfehlungen angesehen werden [13]. Dieses Vorgehen wird auch bei der Auswahl von Quell- und Referenzleitlinien im Rahmen der Erstellung Nationaler VersorgungsLeitlinien (NVL) beschrieben [14].

Domäne 3 umfasst acht Items, Domäne 6 zwei Items. Diese wurden auf einer 7-Punkte Skala von 1 = „trifft überhaupt nicht zu“ bis 7 = „trifft vollständig zu“ bewertet. Die Bewertung wurde von zwei Personen unabhängig durchgeführt. Für jede Domäne wurde ein standardisierter Domänenwert berechnet. Dazu wurden die einzelnen Werte der Items einer Domäne addiert. Die erreichte Gesamtzahl wurde als prozentualer Anteil der maximal möglichen Punktzahl der jeweiligen Domäne dargestellt [12]. Ein Wert > 70% wurde als hohe, ein Wert < 30% als niedrige methodische Qualität und Werte dazwischen als moderate methodische Qualität eingestuft.

## Ergebnisse

Zu 27 IGeL-Themen lagen aktuelle Bewertungen vor. 20 (74%) dieser Themen bezogen sich auf den Bereich der Früherkennung von Erkrankungen. Knapp die Hälfte (n = 12) der betrachteten IGeL enthielten die Bewertung „unklar“. 13 IGeL waren als „tendenziell negativ“, eine IGeL als „negativ“ und eine IGeL als „tendenziell positiv“ eingestuft.

Es wurden 41 LL, die Empfehlungen zu 24 der 27 IGeL-Themen enthielten, in die Analyse eingeschlossen. Über die Hälfte der LL stammte aus Deutschland (n = 11) und den USA (n = 12). Unter den elf deutschen LL handelte es sich bei vier LL um nicht-evidenzbasierte LL (S1- oder S2k-LL). Sechs LL enthielten zu mehr als einem IGeL-Thema Empfehlungen. Die methodische Qualität der evidenzbasierten LL (ohne S1- und S2k-LL), ausgedrückt durch den standardisierten Domänenwert, betrug für Domäne 3 im Mittel 62,3% (Standardabweichung, SD: 18,6%) (Spannweite: 82,3% - 16,7%). Bezogen auf Domäne 6 erreichten alle eingeschlossenen LL (ein-

schließlich S1- und S2k-LL) im Mittel Werte von 70,4% (SD: 22,5%) (Spannweite: 100% - 16,7%). Eine Übersicht über die eingeschlossenen LL pro IGeL-Thema mit den erreichten Domänenwerten findet sich in *Anhang B* (online verfügbar).

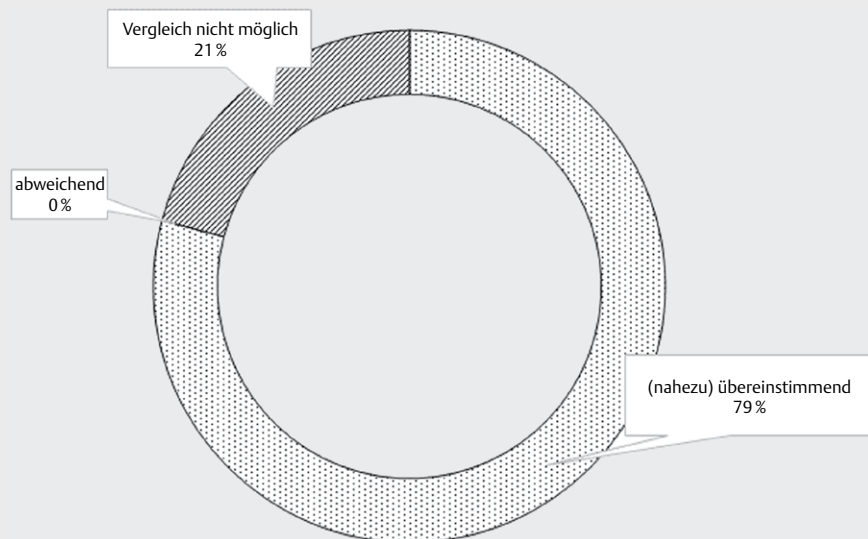
19 Bewertungen (79%) stimmten ganz oder nahezu mit den LL-Empfehlungen überein (► **Abb. 2**). Zu fünf IGeL-Themen (21%) war kein Abgleich möglich. Gründe dafür waren folgende:

- Die Empfehlungen aus unterschiedlichen LL waren inkonsistent (n = 1).
- Die Empfehlungen waren zwar konsistent (gleiche Empfehlungsrichtung), hatten aber eine unterschiedliche Empfehlungsstärke, sodass im Abgleich keine eindeutige Zuordnung zu „(nahezu) übereinstimmend“ oder „abweichend“ möglich war (n = 2).
- Die Empfehlungen waren spezifischer und / oder bezogen sich auf mehrere unterschiedliche Verfahren (n = 2).

Zehn von 13 IGeL, die im IGeL-Monitor (tendenziell) negativ bewertet wurden, wurden auch in den LL nicht empfohlen. Zu einer dieser IGeL-Bewertungen (Screening auf Hirnleistungsstörungen zur Demenzprävention) waren die Empfehlungen inkonsistent. Eine zusammenfassende Übersicht über die Empfehlungen und den Abgleich mit den IGeL-Bewertungen ist in ► **Tab. 1** dargestellt. Eine ausführliche Darstellung findet sich in *Anhang C* (online verfügbar). In *Anhang D* (online verfügbar) sind zudem die jeweiligen Evidenz- und/oder Empfehlungsgraduierungen der eingeschlossenen LL aufgeführt.

## Diskussion

In unserer Analyse stimmten mehr als drei Viertel der IGeL-Bewertungen des IGeL-Monitors, zu denen LL-Empfehlungen identifiziert werden konnten, ganz oder nahezu mit den Empfehlungen überein. Die meisten LL wiesen bezogen auf die Domänen 3 und 6 des



► **Abb. 2** Übersicht Ergebnis Abgleich Leitlinien und IGeL-Bewertungen.

► **Tab. 1** Ergebnisse Abgleich Leitlinienempfehlungen und IGeL-Bewertungen.

IGeL Thema	Bewertung IGeL-Monitor	Anzahl Leitlinien (LL)	Zusammenfassung der Leitlinienempfehlungen	Bewertung des Abgleichs
Akupunktur in der Schwangerschaft	unklar	1	Schwach positive Empfehlung, bezieht sich primär auf die Information zur Akupunktur (Verfahren selber für Vergleich als unklar eingestuft).	(nahezu) übereinstimmend
Akupunktur zur Migräne-Prophylaxe	tendenziell positiv	2	Zwei schwach positive Empfehlungen für eine Akupunktur, wenn eine medikamentöse Therapie nicht geeignet ist und nicht wirkt.	(nahezu) übereinstimmend
Akupunktur bei Spannungskopfschmerz	unklar	1	Schwach positive Empfehlung für eine Akupunktur.	(nahezu) übereinstimmend
EKG zur Früherkennung einer koronaren Herzerkrankung	tendenziell negativ	2	Eine LL spricht sich stark gegen eine Früherkennung aus. Bei einer LL ist unklar, ob es sich bei der Zielpopulation um Personen mit einem erhöhten Risiko für eine koronare Herzkrankheit handelt; zudem fokussiert die Empfehlung auf die Durchführung eines EKGs zur Bewertung der körperlichen Belastbarkeit vor Beginn eines körperlichen Trainingsprogramms (Vergleich für diese LL nicht möglich)	(nahezu) übereinstimmend
Farbstofflaserbehandlung von unkomplizierten frühkindlichen Hämangiomen	tendenziell negativ	1	Schwach positive Empfehlung für operativen Eingriff und Lasertherapie als Therapiemöglichkeiten bei ausgewählten Hämangiomen (Empfehlung spezifischer und mehrere Verfahren adressiert).	Vergleich nicht möglich
Früherkennung von Schilddrüsenfunktionsstörungen mittels TSH-Bestimmung	tendenziell negativ	1	Eine stark negative Empfehlung gegen eine Früherkennung.	(nahezu) übereinstimmend
Glaukom-Früherkennung mittels Augenspiegelung mit Augeninnendruckmessung, OCT oder HRT <sup>®</sup>	tendenziell negativ	1	Statement, dass aufgrund unzureichender Evidenz keine Empfehlung für oder gegen eine Früherkennung möglich ist.	(nahezu) übereinstimmend
HbA1c-Bestimmung zur Früherkennung eines Diabetes	unklar	4	Vier LL empfehlen das Verfahren gleichwertig zu dem Vergleichsverfahren.	(nahezu) übereinstimmend
Laserablation bei Varizen	unklar	1	Positive Empfehlungen für Verfahren der endovenösen thermischen Ablation (mehrere Verfahren adressiert).	Vergleich nicht möglich
Lichttherapie bei Akne	unklar	2	Zwei schwach positive Empfehlungen für eine Lichttherapie.	(nahezu) übereinstimmend
Lungenfunktionstestung mittels Spirometrie bei asymptomatischen Erwachsenen	tendenziell negativ	1	Stark negative Empfehlung gegen eine Früherkennung.	(nahezu) übereinstimmend
M2-PK-Stuhltest zur Früherkennung von Darmkrebs	unklar	1	Schwach negative Empfehlung gegen Früherkennung mittels M2-PK-Stuhltets	(nahezu) übereinstimmend
Magnetresonanztomographie der Brust zur Krebsfrüherkennung	tendenziell negativ	2	Eine LL spricht sich schwach die andere stark gegen eine Früherkennung mittels MRT aus. aus.	(nahezu) übereinstimmend
Magnetresonanztomographie zur Früherkennung einer degenerativen Demenz	tendenziell negativ	1	Stark negative Empfehlung gegen eine Früherkennung.	(nahezu) übereinstimmend
PSA-Test zur Früherkennung von Prostatakrebs	tendenziell negativ	8	8 Leitlinien raten von einem generellen Screening mittels PSA-Test ab (meist schwach negativ) bzw. geben an, dass eine Aufklärung über Nutzen und Schaden der Untersuchung erfolgen soll.	(nahezu) übereinstimmend
Screening auf das Vorliegen einer Verengung der Halsschlagadern mittels Duplexsonographie	tendenziell negativ	3	Drei stark negative Empfehlungen gegen ein Screening.	(nahezu) übereinstimmend
Screening auf Hirnleistungsstörungen zur Demenzprävention (Hirnleistungs-Check)	tendenziell negativ	3	Zwei LL sprechen sich stark gegen eine Früherkennung aus, eine dritte LL gibt aufgrund unzureichender Evidenz keine Empfehlung für oder gegen ein Screening (Inkonsistenz).	Vergleich nicht möglich

► Tab. 1 Fortsetzung

IGeL Thema	Bewertung IGeL-Monitor	Anzahl Leitlinien (LL)	Zusammenfassung der Leitlinienempfehlungen	Bewertung des Abgleichs
Screening auf Vitamin D Mangel	unklar	1	Statement, dass aufgrund unzureichender Evidenz keine Empfehlung für oder gegen ein Screening möglich ist.	(nahezu) übereinstimmend
Toxoplasmose-Test in der Schwangerschaft	unklar	2	Eine LL spricht sich schwach gegen ein allgemeines Screening aus. Eine LL differenziert zwischen verschiedenen Risikogruppen (Vergleich für diese LL nicht möglich)	(nahezu) übereinstimmend
Transrektaler Ultraschall zur Früherkennung eines Prostatakarzinoms	tendenziell negativ	1	Ein evidenzbasiertes Statement mit mittlerer Evidenzkategorie gegen bildgebende Verfahren zur Früherkennung.	(nahezu) übereinstimmend
Transvaginaler Ultraschall zur Früherkennung von Krebserkrankungen des Gebärmutterkörpers	tendenziell negativ	2	Eine LL spricht sich schwach die andere stark gegen eine Früherkennung mittels transvaginalen Ultraschalls aus.	(nahezu) übereinstimmend
Transvaginaler Ultraschall zur Früherkennung eines Ovarialkarzinoms	negativ	4	Vier stark negative Empfehlungen gegen eine Früherkennung.	(nahezu) übereinstimmend
Ultraschall der Brust zur Krebsfrüherkennung	unklar	2	Eine stark und eine schwach negative Empfehlung gegen Verfahren zur Früherkennung bzw. als alleinige Methode der Früherkennung (aufgrund unterschiedlicher Empfehlungsstärken würde sich im Vergleich mit der IGeL-Bewertung einmal eine „(nahezu) Übereinstimmung“, einmal eine „Abweichung“ ergeben).	Vergleich nicht möglich
Vaginaler und rektaler Abstrich auf B-Streptokokken in der Schwangerschaft	unklar	2	Eine stark und eine schwach positive Empfehlung für universelles bzw. risikobasiertes Screening (aufgrund unterschiedlicher Empfehlungsstärken würde sich im Vergleich mit der IGeL-Bewertung einmal eine „(nahezu) Übereinstimmung“, einmal eine „Abweichung“ ergeben).	Vergleich nicht möglich

a: 3 IGeL-Bewertungen (Augenspiegelung mit Augeninnendruckmessung zur Glaukom-Früherkennung, OCT zur Früherkennung eines Glaukoms und Heidelberg Retina Tomographie zur Glaukom-Früherkennung) wurden zu einem IGeL Thema zusammengefasst.

AGREE II-Instrumentes eine moderate bis hohe methodische Qualität auf. Leichte Abweichungen (eine Stufe Unterschied zwischen IGeL-Bewertung und Empfehlungskategorie) zwischen den IGeL-Bewertungen und Empfehlungen bzw. Stärke der Empfehlungen wurden als „nahezu übereinstimmend“ eingestuft. Dieses Vorgehen wurde gewählt, da ein 1:1 Abgleich in der Regel nicht möglich ist. Leichte Abweichungen wurden zum Beispiel für vier IGeL mit „unklarer“ Bewertung festgestellt, für die in den LL schwach positive bzw. schwach negative Empfehlungen gegeben wurden. LL-Empfehlungen werden im Gegensatz zu den IGeL-Bewertungen neben der wissenschaftlichen Evidenz auch von verschiedenen anderen Faktoren beeinflusst, die nicht immer explizit dargelegt und für Ausstehende nachvollziehbar sind, wie beispielsweise gruppendynamische Prozesse, Expert\*innenwissen, Präferenzen von Patient\*innen, Anwendbarkeit und Interessenkonflikte [15–20]. Insbesondere, wenn keine hochwertige Evidenz und Unsicherheiten bei der Nutzen-Schaden-Abwägung vorliegen, fließen häufig andere implizite Faktoren in die Formulierung von Empfehlungen ein [21]. In der vorliegenden Analyse zeigte sich dies beispielsweise bei der IGeL Screening auf Hirnleistungsstörungen zur Demenzprävention, zu der inkonsistente Empfehlungen gefunden wurden. Während die LL der USPSTF aufgrund unzureichender Evidenz keine Empfehlungen für oder gegen ein Screening [22] gab, sprachen sich zwei weitere LL stark dagegen aus [23, 24]. Eine der LL mit star-

ker Negativempfehlung begründete dies aber wie die USPSTF-LL mit fehlender Evidenz [24], bei der anderen LL handelte es sich bei der Empfehlung um einen Expertenkonsens („Good Clinical Practice“) [23].

Zu sieben IGeL mit der Bewertung „tendenziell negativ“ war die Einstufung in den LL zum Teil „strenger“, das heißt, es wurden stark negative Empfehlungen zu den IGeL gegeben. Insgesamt wurden jedoch keine grundlegenden Abweichungen festgestellt.

### IGeL, die mehr schaden als nutzen

Als besonders kritisch sind IGeL anzusehen, die nachgewiesenermaßen mehr schaden als nutzen. Die IGeL „Transvaginaler Ultraschall zur Früherkennung eines Ovarialkarzinoms“ enthielt als einzige IGeL aus unserer Analyse eine negative Bewertung durch das Team des IGeL-Monitors, da sich hierfür keine Hinweise auf einen Nutzen, aber Belege für Schaden fanden. Diese Einschätzung war in unserer Analyse konsistent zu mehreren LL mit starken Negativempfehlungen. In einer Untersuchung von Wegwarth und Gigenzer (2018) wurde festgestellt, dass trotzdem über die Hälfte der Gynäkolog\*innen in den USA ein solches Screening ihren Patientinnen empfiehlt [25]. Die Ärzt\*innen, die ein Screening empfehlen, überschätzten dabei dessen Nutzen und unterschätzten den Schaden häufiger als die Kolleg\*innen, die kein Screening empfehlen [25]. Als häufigste Gründe für das Screening wurden die Erwar-



tungen von Patientinnen angegeben, alles im Kampf gegen Krebs getan zu haben und die Annahme, dass Screening die krankheits-spezifische Mortalität senken kann [26]. Auch in anderen Studien zeigte sich, dass sowohl von Ärztinnen als auch von Patient\*innen der Nutzen medizinischer Interventionen häufig überschätzt und der Schaden unterschätzt wird [27], was ein Einfallstor für das Angebot und die Inanspruchnahme von IGeL darstellt. Auch in Deutschland wird diese IGeL regelhaft angeboten, sie gehört laut IGeL-Report 2020 zu den am zweithäufigsten angebotenen IGeL [3]. Die Gründe hierfür sind nicht klar, wahrscheinlich aber lassen sich einige Gründe aus der Untersuchung in den USA auf Deutschland übertragen.

## Aufklärung über Nutzen und Schaden von IGeL

Insbesondere Untersuchungen zur Krebsfrüherkennung brauchen eine umfassende Aufklärung [28]. Entsprechend sollen auch laut Nationalem Krebsplan Informationsangebote über Nutzen und Risiken der Krebsfrüherkennung mit dem Ziel einer informierten Entscheidung verbessert werden [29]. In unserer Analyse spiegelte sich dies vor allem in einigen Empfehlungen zum PSA-Test zur Früherkennung von Prostatakrebs wider, in denen weniger der Test selber als die Information über Nutzen und Schaden im Fokus steht. Generell erscheint die Beratung zu IGeL in Deutschland verbesserungswürdig zu sein. So waren laut Umfragen nur etwa die Hälfte der Versicherten mit der Aufklärung durch die behandelnde Ärztin / den behandelnden Arzt zufrieden bzw. fühlten sich sehr gut oder gut beraten [2, 3].

## Limitationen

In die Analyse wurden ausschließlich LL eingeschlossen, die in Deutsch oder Englisch publiziert wurden. Dies bedeutet, dass thematisch relevante LL, die in einer anderen Sprache publiziert wurden, möglicherweise fehlen. Weiterhin könnten die eingeschlossenen deutschen nicht-evidenzbasierten und / oder älteren (> 5 Jahre) LL einen verzerrenden Einfluss auf die Ergebnisse der Analyse haben. In unserer Analyse betraf dies fünf IGeL-Themen. Zu all diesen Themen wurde aber mindestens eine weitere LL identifiziert, die evidenzbasiert und < 5 Jahre alt war. Zu drei IGeL waren die Empfehlungen der nicht-evidenzbasierten deutschen LL konsistent mit den Empfehlungen der evidenzbasierten LL. Eine Empfehlung zu vaginalem und rektalem Abstrich auf B-Streptokokken in der Schwangerschaft war zwar konsistent mit der Empfehlung einer evidenzbasierten LL, hatte aber einen stärkeren Empfehlungsgrad. Eine ältere evidenzbasierte deutsche LL aus dem Jahr 2016 enthielt eine Negativ-Empfehlung zu Screening auf Hirnleistungsstörungen zur Demenzprävention, die einerseits inkonsistent zu einer LL aus dem Jahr 2020, andererseits aber konsistent mit der Empfehlung einer weiteren LL aus dem Jahr 2020 war. Insgesamt hatten die eingeschlossenen nicht-evidenzbasierten oder älteren deutschen LL somit eher wenig Einfluss auf das Gesamtergebnis.

## Schlussfolgerung

In der Gesamtschau stimmen die Aussagen aus den IGeL-Bewertungen mit den Empfehlungen aus aktuellen evidenzbasierten LL überein. Insbesondere zu (tendenziell) negativ bewerteten Leistungen, die auch in LL nicht empfohlen, aber trotzdem regelhaft in der ärztlichen Praxis angeboten werden, scheint es wichtig, Versi-

cherte ehrlich und transparent über die Evidenz aufzuklären, um eine informierte Entscheidung für oder gegen eine IGeL zu ermöglichen.

### Merke

- Der Großteil der IGeL-Bewertungen des IGeL-Monitors steht im Einklang mit evidenzbasierten Leitlinienempfehlungen.
- Die meisten IGeL, die im IGeL-Monitor als (tendenziell) negativ eingestuft wurden, werden auch in Leitlinien nicht empfohlen.
- Die Empfehlungen aus unterschiedlichen Leitlinien zu den betrachteten IGeL-Themen sind größtenteils konsistent.
- Eine unsichere oder fehlende Evidenzbasis kann in Leitlinien zu schwach positiven oder schwach negativen Empfehlungen („kann empfohlen werden“ / „kann nicht empfohlen werden“) führen.
- Versicherte sollten ehrlich und transparent über die Evidenz zum Nutzen und Schaden und ggf. auch über die Empfehlungen aus evidenzbasierten Leitlinien informiert werden, um eine informierte Entscheidung für oder gegen eine IGeL zu ermöglichen.

## Interessenkonflikt

Die Autorinnen sind Mitarbeiterinnen des Medizinischen Dienst Bund. Dieser ist Herausgeber des IGeL-Monitors. Der Medizinische Dienst Bund ist eine Körperschaft des öffentlichen Rechts und wird von den Medizinischen Diensten der Länder finanziert.

## Literatur

- [1] Bundesministerium für Gesundheit. Was sind Individuelle Gesundheitsleistungen (IGeL)? (2016). Im Internet: <https://www.bundesgesundheitsministerium.de/service/begriffe-von-a-z/i/igel.html>; Stand: 24.01.2023
- [2] Zok K. Private Zusatzleistungen in der Arztpraxis. Ergebnisse einer bundesweiten Repräsentativ-Umfrage unter gesetzlich Versicherten (01/2019) Im Internet <https://www.wido.de/publikationen-produkte/widomonitor/widomonitor-1-2019/>; Stand: 23.01.2023
- [3] Medizinischer Dienst des Spitzenverbandes Bund der Krankenkassen (MDS). IGeL-Report 2020. Im Internet: [https://www.igel-monitor.de/fileadmin/Downloads/Presse/2020\\_08\\_25\\_PK\\_IGEL\\_Report\\_2020\\_COVID/2020\\_08\\_25\\_Ausfuehrlich\\_IGeL\\_Report\\_2020.pdf](https://www.igel-monitor.de/fileadmin/Downloads/Presse/2020_08_25_PK_IGEL_Report_2020_COVID/2020_08_25_Ausfuehrlich_IGeL_Report_2020.pdf); Stand: 23.01.2023
- [4] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)-Ständige Kommission Leitlinien. AWMF-Regelwerk „Leitlinien“. 2. Auflage (2020). Im Internet: <http://www.awmf.org/leitlinien/awmf-regelwerk.html>; Stand: 23.01.2023. In: 2020
- [5] Institute of Medicine (IOM). Clinical Practice Guidelines We Can Trust. Washington (DC): The National Academies Press; 2011. DOI: 10.17226/13058
- [6] Härter M, Koch-Gromus U. Qualitätsentwicklung im Gesundheitswesen: Sind wir schon so weit, wie wir wollten? Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2022; 65: 267–269. DOI: 10.1007/s00103-022-03503-4

- [7] Nothacker M, Muche-Borowski C, Kopp IB. 20 Jahre ärztliche Leitlinien in Deutschland – was haben sie bewirkt? *Z Evid Fortbild Qual Gesundheitswes* 2014; 108: 550–559. DOI: 10.1016/j.zefq.2014.10.012
- [8] Holleczeck B, Brenner H. Provision of breast cancer care and survival in Germany - results from a population-based high resolution study from Saarland. *BMC Cancer* 2014; 14: 757. DOI: 10.1186/1471-2407-14-757
- [9] Wöckel A, Kurzeder C, Geyer V et al. Effects of guideline adherence in primary breast cancer – a 5-year multi-center cohort study of 3976 patients. *Breast* 2010; 19: 120–127. DOI: 10.1016/j.breast.2009.12.006
- [10] A&W Online. IGeL-Monitor: SpiFa wehrt sich gegen "mediale Generalurteilung" von IGeL-Leistungen (09.09.2019). Im Internet: <https://www.arzt-wirtschaft.de/finanzen/honorare/igel-monitor-spifa-wehrt-sich-gegen-mediale-generalurteilung-von-igel-leistungen/>; Stand: 23.01.2023
- [11] Ärzteblatt. IGeL-Leistungen laut MDS nicht in Einklang mit medizinischen Empfehlungen (03.05.2018). Im Internet: <https://www.aerzteblatt.de/nachrichten/94893/IGeL-Leistungen-laut-MDS-nicht-in-Einklang-mit-medizinischen-Empfehlungen>; Stand: 23.01.2023
- [12] AGREE Next Steps Consortium. Appraisal of guidelines for research & evaluation II: AGREE II instrument (12/2017). Im Internet: <https://www.agreertrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf>; Stand: 23.01.2023. In
- [13] Hoffmann-Eßer W, Siering U, Neugebauer EAM et al. Guideline appraisal with AGREE II: online survey of the potential influence of AGREE II items on overall assessment of guideline quality and recommendation for use. *BMC Health Serv Res* 2018; 18: 143. DOI: 10.1186/s12913-018-2954-8
- [14] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Programm für Nationale VersorgungsLeitlinien – Methodenreport, 5. Auflage. Version 1 (2017). Im Internet: [www.leitlinien.de/methodik](http://www.leitlinien.de/methodik); Stand: 23.01.2023. DOI: 10.6101/AZQ/000169
- [15] Djulbegovic B, Reljic T, Elqayam S et al. Structured decision-making drives guidelines panels' recommendations "for" but not "against" health interventions. *J Clin Epidemiol* 2019; 110: 23–33. DOI: 10.1016/j.jclinepi.2019.02.009
- [16] Nejtgaard CH, Bero L, Hróbjartsson A et al. Conflicts of interest in clinical guidelines, advisory committee reports, opinion pieces, and narrative reviews: associations with recommendations. *Cochrane Database Syst Rev* 2020; 12: Mr000040. DOI: 10.1002/14651858.MR000040.pub3
- [17] Norris SL, Burda BU, Holmer HK et al. Author's specialty and conflicts of interest contribute to conflicting guidelines for screening mammography. *J Clin Epidemiol* 2012; 65: 725–733. DOI: 10.1016/j.jclinepi.2011.12.011
- [18] Pagliari C, Grimshaw J. Impact of group structure and process on multidisciplinary evidence-based guideline development: an observational study. *J Eval Clin Pract* 2002; 8: 145–153. DOI: 10.1046/j.1365-2753.2002.00333.x
- [19] Raine R, Sanderson C, Hutchings A et al. An experimental study of determinants of group judgments in clinical guideline development. *Lancet* 2004; 364: 429–437. DOI: 10.1016/s0140-6736(04)16766-4
- [20] Richter Sundberg L, Garvare R, Nyström ME. Reaching beyond the review of research evidence: a qualitative study of decision making during the development of clinical practice guidelines for disease prevention in healthcare. *BMC Health Serv Res* 2017; 17: 344. DOI: 10.1186/s12913-017-2277-1
- [21] Djulbegovic B, Hozo I, Li SA et al. Certainty of evidence and intervention's benefits and harms are key determinants of guidelines' recommendations. *J Clin Epidemiol* 2021; 136: 1–9. DOI: 10.1016/j.jclinepi.2021.02.025
- [22] Owens DK, Davidson KW, Krist AH et al. Screening for Cognitive Impairment in Older Adults: US Preventive Services Task Force Recommendation Statement. *Jama* 2020; 323: 757–763. DOI: 10.1001/jama.2020.0435
- [23] Deutsche Gesellschaft für Psychiatrie und Psychotherapie PuN, Deutsche Gesellschaft für Neurologie. S3-Leitlinie "Demenzen". AWMF-Registernummer: 038-013 (2016). Im Internet: <https://www.awmf.org/leitlinien/detail/II/038-013.html>; Stand 23.01.2023
- [24] Ismail Z, Black SE, Camicioli R et al. Recommendations of the 5th Canadian Consensus Conference on the diagnosis and treatment of dementia. *Alzheimers Dement* 2020; 16: 1182–1195. DOI: 10.1002/alz.12105
- [25] Wegwarth O, Gigerenzer G. US gynecologists' estimates and beliefs regarding ovarian cancer screening's effectiveness 5 years after release of the PLCO evidence. *Sci Rep* 2018; 8: 17181. DOI: 10.1038/s41598-018-35585-z
- [26] Wegwarth O, Pashayan N. When evidence says no: gynaecologists' reasons for (not) recommending ineffective ovarian cancer screening. *BMJ Qual Saf* 2020; 29: 521–524. DOI: 10.1136/bmjqs-2019-009854
- [27] Hoffmann TC, Del Mar C. Clinicians' Expectations of the Benefits and Harms of Treatments, Screening, and Tests: A Systematic Review. *JAMA Intern Med* 2017; 177: 407–419. DOI: 10.1001/jamainternmed.2016.8254
- [28] Koch K, Thomas S, Windeler J. IGeL oder Nicht-IGeL. Krebsfrüherkennung braucht immer eine umfassende Aufklärung. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2014; 57: 334–342. DOI: 10.1007/s00103-013-1909-9
- [29] Bundesministerium für Gesundheit. Ziele des Nationalen Krebsplans (2008). Im Internet: <https://www.bundesgesundheitsministerium.de/themen/praevention/nationaler-krebsplan/handlungsfelder/ziele-des-nationalen-krebsplans.html>; Stand: 23.01.2023