

# It's MAGIC – einfaches geriatrisches Basisassessment für die Hausarztpraxis validiert

*It's MAGIC –*

*Validation of a Simple Geriatric Assessment in Family Medicine*

Ulrike Junius-Walker<sup>1</sup>, Nick Daether-Kracke<sup>1</sup>, Olaf Krause<sup>1,2</sup>

**Hintergrund:** In der Praxis ist es häufig schwierig, ältere vulnerable Patienten zu identifizieren. Dazu kann ein hausärztlich geriatrisches Assessment beitragen. Ausgehend von dem europäischen STEP-Assessment, wurde eine deutlich kürzere Version „MAGIC“ (manageable geriatric assessment) entwickelt. In einem Praxistest sollten nun Erkenntnisse zur Testgüte des MAGIC geliefert werden.

**Methoden:** In einer Querschnittstudie mit zweizeitiger Erhebung erhielten hausärztliche Patienten ab 70 Jahren zunächst das MAGIC-Assessment und eine Woche später das längere STEP-Assessment. Jeder vierte Patient nahm zusätzlich an einem kognitiven Interview teil, in dem er zum Verständnis und zur Intention der Fragestellungen des MAGIC Auskunft gab. Zur Analyse der Testgüte wurde die Test-Retest-Reliabilität (Übereinstimmungsrate und Spearman-Korrelation) bestimmt. Die Auswertung der kognitiven Interviews erfolgte qualitativ mittels einer sogenannten thematischen Strukturierung mit induktiver Kategorienbildung nach Mayring.

**Ergebnisse:** 30 Hausärzte aus dem südlichen Raum Hannovers nahmen an der Studie teil. Aus diesen Praxen wurden insgesamt 60 Patienten mit einem Altersdurchschnitt von 79,9 Jahren rekrutiert (Teilnahmerate 47 %). Das MAGIC-Assessment deckte im Mittel vier von neun abgefragten Gesundheitsproblemen auf. Patienten verstanden die MAGIC-Fragen mit wenigen Ausnahmen. Sie bezogen die Fragen einerseits auf die medizinisch-körperliche Funktionsfähigkeit und andererseits auf konkrete Alltagsauswirkungen. Die Test-Retest-Übereinstimmungsrate von MAGIC lag zwischen 83,3 % für das Item Depression und 100 % für Angaben zur sozialen Teilhabe.

**Schlussfolgerungen:** MAGIC zur Identifizierung älterer vulnerabler Patienten hat im Praxistest zeitlich stabile Testergebnisse und ein gutes Frageverständnis erbracht. Die MAGIC-Fragen veranlassen Patienten dazu, eine Pro-

**Background:** It is challenging for family practitioners to identify older vulnerable patients in routine care. A geriatric assessment can contribute to this task. Based on the European STEP-assessment a notably shorter version called "MAGIC" (manageable geriatric assessment) was developed. The aim of this paper is to report on the validity and reliability of MAGIC as determined in a field test.

**Methods:** In a two-stage cross-sectional survey family practice patients (70+ years) received the MAGIC-assessment and one week later the longer version STEP. Additionally, every fourth patient participated in a cognitive interview and provided feedback on the intention and meaning of the MAGIC-questions. To analyse the test-retest-reliability, the MAGIC findings at t0 and t1 were compared and a Spearman correlation performed. Cognitive interviews were assessed using an inductive approach with simple coding frames according to Mayring.

**Results:** 30 family physicians from the southern region of Hannover took part. From these practices a total of 60 patients with an average age of 79.7 years were recruited (participation rate 47 %). MAGIC detected a median of four out of a maximum of nine health problems. In general, patients easily understood the meaning and intention of the MAGIC-questions. In their explanations, they referred to the medical context and also described the consequences for their daily life. Test-retest agreement of MAGIC ranged between 83.3 % for the item "depression" and 100 % for "social participation".

**Conclusions:** MAGIC, to identify vulnerable older patients, has produced stable results and a good understanding of its content. The MAGIC-questions encourage patients to assess their health and personal problems in the context of their everyday life, which is conducive to identifying vulnerable patients' needs. Our results add to the existing knowledge of validity and reliability, and provide an evidence base that is legally required for perform-

<sup>1</sup> Institut für Allgemeinmedizin, Medizinische Hochschule Hannover

<sup>2</sup> Zentrum für Medizin im Alter, Diakoniekrankenhaus Henriettenstiftung, Hannover

Peer reviewed article eingereicht: 04.01.2016, akzeptiert: 26.01.2016

DOI 10.3238/zfa.2016.0169-0175

blembewertung für ihren persönlichen Alltagskontext vorzunehmen, was für eine Identifizierung vulnerabler Patienten und therapeutische Bedarfsanalyse wichtig ist. Unsere Ergebnisse in Ergänzung zu bestehenden Erkenntnissen der Testgüte bilden eine differenzierte Evidenzbasis, wie es für die Nutzung eines geriatrischen Basisassessments gefordert wird.

*Schlüsselwörter: Allgemeinmedizin; ältere Patienten; geriatrisches Basisassessment; Validität; Reliabilität*

ing a geriatric assessment in German family practices.

*Keywords: Family Medicine; Older Patients; Geriatric Assessment; Validity; Reliability*

## Einleitung

Hausärzte übernehmen eine Schlüsselrolle in der ganzheitlichen Betreuung älterer Patienten. Dabei gilt es, die Gefährdungslage einer gesundheitlichen Verschlechterung im Alter gut einzuschätzen. Das ist nicht immer leicht, zumal die Gruppe der selbstständig lebenden älteren Menschen einen sehr heterogenen Gesundheitszustand aufweist und das Alter allein zur Beurteilung geistiger und körperlicher Fähigkeiten kaum geeignet ist. So weisen nach einer Bevölkerungsumfrage zwei Drittel der 70- bis 80-Jährigen Einschränkungen ihrer Alltagsaktivitäten auf, ein Drittel ist jedoch fit [1]. Ein geriatrisches Assessment soll Hausärzte darin unterstützen, diese Gefährdungslage einzuschätzen, um frühzeitig Maßnahmen einzuleiten, die einem Nachlassen von körperlichen, funktionellen und psychosozialen Fähigkeiten entgegenwirken [2].

Bisher gibt es kaum geeignete hausärztliche Assessmentinstrumente im deutschsprachigen Raum zur strukturierten Erhebung von relevanten Altersproblemen. Zu nennen ist jedoch das geriatrische STEP-Assessment, das eigen von einem Verbund europäischer Hausärzte konzipiert wurde. Mit spezifischen Fragen, Performancetests und Laborwerten werden 44 Gesundheitsprobleme erfasst, die aufgrund evidenzbasierter Erkenntnisse eine hohe Relevanz im Alter aufweisen und gleichzeitig ohne viel Aufheben diagnostizierbar bzw. effektiv therapierbar sind. In mehreren Studien wurde STEP erfolgreich eingesetzt, so in Österreich [3], Deutschland [4] und in Südtirol [5]. Die beteiligten Hausärzte schätzten die Detektion bisher unbekannter Gesundheitsprobleme als besonders hilfreich ein und konnten

diese Probleme in einem hohen Maß erfolgreich therapieren [4]. Allerdings zeigte sich auch, dass das STEP-Assessment zeitaufwendig und schwer in den Praxisalltag umzusetzen ist [6, 7].

Ein hausärztliches Assessment muss allerdings in der Breitenversorgung ohne großen Aufwand in der Praxis durchführbar sein. Daher wurde nun das kurze, geriatrische Assessment „MAGIC“ (manageable geriatric assessment), ein Messinstrument mit nur neun Items, auf der Basis von STEP entwickelt. Dies erfolgte in einem partizipativen Forschungsansatz praktischer tätiger Hausärzte. Evidenz- und erfahrungsgestützt priorisierten sie die Gesundheitsthemen, die im Praxisalltag eher zu kurz kommen und die einen hohen Wert für die Selbstständigkeit und Lebensqualität im Alter haben [8].

Nun gilt es, Aussagen über die Testgüte im Rahmen eines Praxistests zu erhalten. Ein gutes Messinstrument sollte valide und reliabel sein. *Reliabilität* beschreibt die Stabilität der gemessenen Werte bei verschiedenen Testdurchführenden bzw. über einen Zeitraum hinweg. Die *Validität* ist ein Maß dafür, inwieweit die in der Messung erzeugten Daten auch tatsächlich die zu messende Größe repräsentieren. Validität lässt sich hinsichtlich Konstrukt-, Kriteriums- und Inhaltsvalidität beurteilen. Ein *konstruktvalider Test* liegt vor, wenn empirisch überprüfbare Aussagen über Zusammenhänge des zugrunde liegenden Konstrukts mit anderen Variablen hergeleitet werden können. *Kriteriumsvalidität* untersucht, inwieweit das Testergebnis mit anderen, praktisch relevanten Kriterien (sogenannten Außenkriterien) übereinstimmt. *Inhaltsvalidität* bezieht sich auf die Eignung und Vollständigkeit des Messinstruments im Hinblick auf das Konstrukt [9]. Zur Überprü-

fung der Eignung gehört insbesondere das Fragenverständnis [10].

Auch das MAGIC-Instrument muss sich einer solchen Prüfung stellen, zumal der Gemeinsame Bundesausschuss im EBM O3360 für ein Assessmentinstrument eine ausreichende Validität gefordert hat [11]. Ziel der vorliegenden Arbeit ist es daher, Validität und Reliabilität des MAGIC-Assessments in einem Praxistest zu überprüfen, wobei bereits veröffentlichte Erkenntnisse zur Testgüte berücksichtigt werden.

## Methoden

### Literaturanalyse

Zur Klärung bereits vorhandener Daten zur Validität und Reliabilität wurde eine pragmatische Literatursuche mit Knowledge Finder ohne Zeitlimitation durchgeführt. Zusätzlicher Fundort war die „Bestandsliteratur“ aus der STEP- und MAGIC-Entwicklung.

### Rekrutierung

Der Praxistest lief als Querschnittstudie mit zweizeitiger Erhebung, die in 2014 in 17 Praxen mit 30 Hausärzten und 60 älteren Patienten im südlichen Raum Hannovers durchgeführt wurde. Insgesamt wurden alle Praxen dieser Region, die schon einmal eine Studie mit dem Institut durchgeführt hatten, angesprochen und weitere durch persönlichen Kontakt, bis die anvisierte Rekrutierungszahl erfüllt war. In jeder Praxis sprach eine Medizinische Fachangestellte an einem festgelegten Rekrutierungstag konsekutiv die Patienten an, die 70 Jahre und älter waren. Ausgeschlossen wurden Patienten, wenn sie kein Telefon hatten, eine be-

Item des MAGIC	Kriteriumsvalidität	Test-Retest-Reliabilität (TRT)
<b>Leistungsfähigkeit</b> Schwierigkeiten alltägliche Arbeiten innerhalb und außerhalb des Hauses zu erledigen [15]	Korrelationen mit dem Subtest „Funktion“ der Short Form 36 (SF-36) $r = 0,62$ [15]	TRT (3 Wochen): $r = 0,67-0,82$ Kappa = $0,36-0,72$ [15]
<b>Sehen und Hören</b> – Probleme Zeitungsschrift zu erkennen – Probleme, Personen auf der anderen Straßenseite zu erkennen – Schwierigkeiten, Gespräche zu verstehen [16]	Vergleich mit Functional Limitations Profile (FLP) Sens. 88 % Spez. 72 %, Korrelation $r = 0,79$ [16]	TRT (3–6 Monate): 29 % Patienten wiesen bei insgesamt 30 Fragen zumindest eine Veränderung auf [17]
<b>Stürze</b> 1 bzw 2 oder mehr Stürze in letzten 6 Monaten? [18]	OR 4,9 für Sturz in nächsten 2 Jahren [19]	TRT für Personen im Betreuten Wohnen Kappa = $0,96$ [20]
<b>Harninkontinenz</b> – Tropfen aus der Blase – Toilette nicht mehr rechtzeitig erreichen [21]	Vergleich zur ärztl. Diagnose: Stress: Sens. 0,59, Spez. 0,92 Drang: Sens. 0,62, Spez. 0,95 Gemischt: Sens. 0,79, Spez. 0,62 [21]	Keine TRT bekannt
<b>Stimmung</b> – niedergeschlagen, deprimiert oder hoffnungslos – weniger Interesse oder Freude [22]	Vergleich zur psychiatrischen Diagnose anhand DSM-IV Sens.: 0,94, Spez. 0,63 bei zumindest einer positiven Antwort [22]*	Keine TRT bekannt
<b>Soziales Umfeld</b> Haben Sie jemanden, – der sich im Notfall um Sie kümmert – auf den Sie sich verlassen und dem Sie sich anvertrauen können? [23]	Erste Fragestellung wurde verändert von „jdn, der Ihnen hilft“ zu „jdn, der sich im Notfall um Sie kümmert“. Keine Kriteriumsvalidität hierzu bekannt [24]	TRT (5 Wochen) Übereinstimmung über alle OARS Items: 91 % [25]
<b>Impfung</b> – Grippe (Influenza) – Wundstarrkrampf (Tetanus) – Diphtherie	Influenza: Korrekte Angabe in 96 % [26] Tetanus: Sens. 0,92, Spez. 0,11; Kappa = $0,03$ [27]	Keine TRT bekannt
<b>Uhrentest</b> [28]	Vergleich zu MMSE mit cut-off 17: Sens. 0,77, Spez. 0,87 [29]*	TRT $0,7-0,94^*$
* Insbesondere für den Fragebogen nach Whooley [22] und für den Uhrentest [29] liegen noch weitere Ergebnisse zur Validität bzw. Reliabilität vor. Sens. = Sensitivität; Spez. = Spezifität		

**Tabelle 1** Bekannte Testgütekriterien für die einzelnen MAGIC-Items

kannte schwere Demenz oder starke Schwerhörigkeit aufwiesen. Alle Teilnehmer wurden über die Studie schriftlich informiert und gaben ihr Einverständnis (Ethikvotum der Medizinischen Hochschule Hannover liegt vor, Nr. 2233–2014). Sobald zwei Patienten pro beteiligtem Arzt eingeschlossen waren, war die Rekrutierung in einer Praxis abgeschlossen.

#### Ablauf

In den jeweiligen hausärztlichen Praxen führte N.DK mit je zwei Patienten nacheinander das MAGIC-Assessment durch. Das Assessment selbst dauerte ca. 10 Minuten. Testauffälligkeiten wurden auf einen Ergebnisbogen übertragen. Die Patienten bewerteten daraufhin die subjektive Relevanz dieser Auffälligkeiten. Sie beantworteten weiterhin standardisierte und offene Fragen zur Machbarkeit und Akzeptanz. Jeder vierte Patient

erhielt ein sogenanntes Probing (ausführlich s.u.). Genau eine Woche später führte N. D.-K. mit den Studienpatienten in einem Hausbesuch das ca. 50-minütige STEP-Assessment durch. Im STEP sind alle neun MAGIC-Items enthalten. Wiederum bewerteten die Patienten die subjektive Relevanz der zutage getretenen Auffälligkeiten. Diesmal beantworteten sie standardisierte und offene Fragen zur Machbarkeit und Akzeptanz des STEP-Assessments.

Schließlich wurden die MAGIC- und STEP-Ergebnisse eines jeden Patienten zur Information und Bewertung den Hausärzten vorgelegt.

#### Probing

Probing als Erhebungsmethode zum Frageverständnis des MAGIC-Assessments: Das Probing ist eine Methode des kognitiven Interviews zum Frageverständnis eines Testinstruments [12]. Das

Frageverständnis wiederum gibt Hinweise über die Inhaltsvalidität der neun Items des MAGIC. Beim retrospektiven Probing werden den Interviewten nach der Testdurchführung festgelegte Fragen gestellt, wobei spontan ergänzende Fragen folgen können [13]. Für unsere Studie lauteten die Standardfragen: „Woran genau dachten Sie bei der Frage zu ...?“ und „Was sollte Ihrer Meinung nach mit der Frage bezweckt werden?“.

#### Auswertung

Die Antworten zu den offenen Fragen der 14 Teilnehmer, die beim Probing teilnahmen, wurden schriftlich festgehalten. Es folgte eine sogenannte thematische Strukturierung mit induktiver Kategorienbildung nach Mayring [14]. Zur Bestimmung der Test-Retest-Reliabilität der MAGIC-Items zum Zeitpunkt t0 (Hausarztpraxis) und t1 (1 Woche später als Hausbesuch) wurde die beobachtete

Übereinstimmung der patientenseitigen Antworten für jedes Item ermittelt und Korrelationen vorgenommen.

## Ergebnisse

### Literaturanalyse

Für 5 der 9 MAGIC-Fragen liegen Aussagen zur Test-Retest-Reliabilität vor (Sehen und Hören als Teil eines umfassenderen Fragebogens, Stürze, Uhrentest und Leistungsfähigkeit, Tab. 1). Die Stabilität der Antworten ist für Stürze und das soziale Umfeld als hoch zu bewerten; für Leistungsfähigkeit, Sehen und Hören sowie den Uhrentest ist sie mäßig bis hoch. Für 8 der 9 Magic-Fragen liegen Aussagen zur Validität vor, wobei es sich hier durchgehend um Untersuchungen zur Kriteriumsvalidität handelt (Tab. 1). Eine Untersuchung zum Frageverständnis ist unserer Kenntnis nach noch nicht vorgenommen worden.

### Studienteilnehmer

30 von insgesamt 70 angesprochenen Hausärzten nahmen an der Studie teil (Teilnahmerate 40 %). Sie waren im Schnitt 50,9 Jahre ( $\pm 9,5$ ) alt und überwiegend männlich (73 %). 29 Ärzte waren eigenverantwortlich als Hausärzte tätig: 15 als Fachärzte für Allgemeinmedizin, 14 als hausärztliche Internisten. Ein Arzt arbeitete als Arzt in Weiterbildung Allgemeinmedizin.

In den Praxen wurden 129 Patienten angesprochen, 60 erklärten sich zur Teilnahme bereit (47 %). Das durchschnittliche Patientenalter belief sich auf 79,7 Jahre ( $\pm 4,7$ ). 35 Patienten waren weiblich (58 %). Die Patienten hatten überwiegend eine deutsche Staatsangehörigkeit (98 %), lebten in einem Privathaushalt (98 %) und waren gesetzlich krankenversichert (95 %). 23 Patienten (38 %) lebten allein. Im MAGIC wiesen die Patienten im Median vier Gesundheitsprobleme (2–8) auf. Fast die Hälfte (46,7 %) nahm sechs und mehr Medikamente ein. 12 % hatten eine Pflegestufe erhalten.

### Frageverständnis zum MAGIC

Jeder vierte Patient ( $N = 15$ ) nahm an dem Probing teil. Es waren mehr Frauen (71,4 %) und sie waren im Durchschnitt auch etwas älter als die Gesamtstichpro-

be (81,3 Jahre  $\pm 5,1$ ). Es galt für die Items zu prüfen, was erstens genau unter der Fragestellung verstanden wird, und zweitens, welche Absicht mit der Frage verfolgt wird. Für die letzten beiden Items (Frage zum Impfstatus und Uhrentest) boten die Patienten nur Erklärungen zur mutmaßlichen Absicht und nicht zum Verständnis an.

Tabelle 2 besteht aus zwei Spalten. In der ersten Spalte sind die inhaltlichen Kategorien aufgeführt, die das Frageverständnis eines Items ausmachen. In der zweiten Spalte befinden sich die Kategorien zur Frageintention. Für jede Kategorie ist zusätzlich die Anzahl der Patienten angegeben, die Aussagen hierzu gemacht haben. Ergänzend finden sich beispielhaft Zitate.

Beim Frageverständnis der Items bezogen die Patienten die Inhalte oft auf ihre Alltagssituationen bzw. -erlebnisse; weniger häufig äußerten sie sich zu rein medizinischen Aspekten. Die Assoziationen waren entsprechend breit, z.B. beim Frageverständnis zur „Leistungsfähigkeit: alltägliche Arbeiten“ reichten die Patientenvorstellungen von der Fähigkeit sich im Haus zu bewegen bis zur beruflichen Arbeit außer Haus. Beim Thema Harninkontinenz beispielsweise stellten einige Patienten das ungewollte Wasserlassen in den Vordergrund, andere die Belastung durch Schamgefühle.

Zu den Intentionen der Fragestellungen gaben die Patienten wiederum medizinische und alltägliche Themen an. Sie benannten als Zweck eine medizinisch diagnostische Abklärung bzw. das Ermitteln des ärztlichen Handlungsbedarfes genauso wie ein Sondieren der eigenen Leistungsfähigkeit und des Hilfsbedarfs im Alltag. Diese doppelte Betrachtungsweise zeigte sich für alle Items – vom Impfen bis zur Harninkontinenz.

Insgesamt gaben die Patienten 225 Antworten hinsichtlich Frageverständnis und Zweck. Nur eine Antwort zum Verständnis „Sehen“ und zwei Antworten zum Zweck „Uhrentest“ interpretierten wir als patientenseitig „missverstehen“ (s. \* in Tab. 2).

### Stabilität der MAGIC-Items

In Abbildung 1 ist dargestellt, wie stabil die teilnehmenden Patienten die neun MAGIC-Fragen, die auch im STEP-Assessment enthalten waren, im Ein-Wochenabstand beantworteten. Die Test-Retest-

Übereinstimmungsrate lag zwischen 83,3 % für das Item Depression und 100 % für Angaben zum sozialen Umfeld. Die Berechnung der Korrelationen belegte, dass ein sehr hoher Zusammenhang zwischen t0 und t1 bei den Antworten zur Grippeimpfung und zum sozialen Umfeld bestand (Spearman's rho = 0,93 bzw. 1). Für Sturz und Inkontinenz gab es eine hohe Korrelation (rho = 0,78 bzw. 0,83) und für die Leistungsfähigkeit Sehen, Hören, Diphtherieimpfung und Depression eine mittlere (rho = 0,67 bzw. 0,58 bzw. 0,68 bzw. 0,65 bzw. 0,59). Eine eher geringe Korrelation war bei der Tetanusimpfung zu verzeichnen (rho = 0,43).

## Diskussion

Die MAGIC-Fragen können größtenteils konstant über eine Woche hinweg beantwortet werden. Patienten verstehen die Inhalte der Fragestellungen, abgesehen von wenigen Ausnahmen, wie medizinisch „intendiert“, verorten sie aber zusätzlich auch im Alltagserleben.

Auffällig ist, dass die Test-Retest-Reliabilität allgemein gebräuchlicher anamnestischer Fragen, wie sie ja nicht nur im MAGIC verwendet werden, bisher kaum untersucht worden ist. Bei unseren hausärztlichen Patienten zeigt sich für die einzelnen Items eine hohe Antwortstabilität mit einer 80- bis 100-prozentigen Übereinstimmung der einzelnen Testergebnisse über den Zeitraum von einer Woche. Beim Uhrentest sahen wir von einer zweiten Testung ab, weil sie häufig eine Belastung für Patienten darstellt [30], ohne dass hier die Aufdeckung einer kognitiven Dysfunktion primäres Untersuchungsziel gewesen wäre. Zusätzlich hatten wir aus früheren Untersuchungen gelernt, dass ein gewisser Teil der positiv Getesteten von sich aus mit dem regelmäßigen Üben der Uhrentests begann. Für den Uhrentest liegen bisher keine Ergebnisse zur Test-Retest-Reliabilität für hausärztlich ambulante Patienten vor, wohl aber für Patienten im spezialisierten Setting und Teilnehmer einer Kohortenstudie [31].

Zum Frageverständnis (Verstehen und vermutete Intention der Fragen) lässt sich feststellen, dass die Patienten neben einer medizinischen Sichtweise auch die persönlichen Konsequenzen eines Gesundheitsproblems im Alltag und Sozialleben abschätzen. Wenn Patienten

Antworten zum Frageverständnis			Antworten zur Frageintention		
Kategorie	N	Beispiel	Kategorie	N	Beispiel
<b>Leistungsfähigkeit</b>					
Haus- und Gartenarbeit	11	„Mitarbeit im Betrieb“ (54) „Bewegung zu Hause, problemlose Toilettengänge“ (35)	Selbständigkeit prüfen	9	„Ist Patient noch in der Lage, den Haushalt zu versorgen?“ (07)
Hobbies	1		Leistungsfähigkeit prüfen	4	
Mobilität	1		Gesellschaftliche Teilhabe prüfen	1	
<b>Sehen</b>					
Sehen allgemein, Fern- oder Nahsicht	13	„Ob das Zeitungslesen Schwierigkeiten bereitet; ob der Nachbarsjunge noch auf dem Spielplatz zu erkennen ist“ (45)	Sehfähigkeit prüfen	5	„Ist augenärztliche Kontrolle notwendig?“ (47)
Lesen mit Sehhilfe	2		Hilfsbedarf erkennen	5	
Lesen bei schlechtem Licht	2		Gesell. Teilhabe ermöglichen	3	
Literaturinteresse	1*		Gefahren erkennen	1	
<b>Hören</b>					
Hören allgemein	1	„Auf beiden Ohren noch gute Hörfähigkeit?“ (54)	Gesell. Teilhabe ermöglichen	6	„Ob man Gefahrensituationen noch akustisch wahrnehmen kann“ (47)
Hören im Zwiegespräch	4		Hörfähigkeit prüfen	5	
Hören in Gesellschaft	12		Hilfsbedarf erkennen	3	
			Gefahren erkennen	1	
<b>Stütze</b>					
Stolpern allgemein	4	„Ob man an der Teppichkante hängen bleibt und dann stolpert“ (54) „... an Schwindelanfälle, die dann zu Zusammenbrüchen führen“ (24)	Ärztlicher Handlungsbedarf	5	„Arzt soll zur Reaktion angehalten werden“ (03), „Sind Hindernisse im Haus vorhanden?“ (35)
Stolpern außer Haus	4		Gefahr stationärer Aufnahme	4	
Stolpern im Haus	3		Sturzgefahr testen	3	
Sturz vom Fahrrad	3		Häuslichkeit checken	2	
Schwindel oder Ohnmacht	3		Koordinationscheck	1	
<b>Harninkontinenz</b>					
Im Alltag	5	„Kann Pat. im Alltag den Harn halten?“ (13), „Ist der Alltag durch Blasenschwäche belastet, z.B. durch Peinlichkeitsempfinden?“ (24)	Diagnostische Abklärung	5	„Inkontinenz-Austestung“ (07), „Sozialleben möglich?“ (24)
Drang und oder Stress	4		Klärung sozialer Konsequenzen	5	
Nach Trinken	4		Inkontinenz testen	3	
Scham	3		Handlungsbedarf erkennen	1	
Nach OP	1				
<b>Stimmung</b>					
Depressive Symptome	11	„Ist man schlecht drauf aufgrund von Krankheit oder Kummer?“ (49), „Hat man noch Lust, Leute zu treffen und sich zu unterhalten?“ (53)	Depression abklären	10	„Depressionsgefährdung“ (20), „Besteht Handlungsbedarf ob der seelischen Situation?“ (13)
Soziale Unterstützung	3		Unterstützungsbedarf klären	4	
Gesellschaftliche Isolierung	3		Geistige Gesundheit testen	2	
<b>Impfung</b>					
			Eigenverantwortung	8	„Ob man gut geschützt ist und die Tipps des Doktors umsetzt“ (24)
			Klärung Schutz	5	
			Therapietreue	1	
<b>Uhrentest</b>					
			Check geistiger Leistungsfähigkeit	12	„Ob Patient noch weiß, wie eine Uhr aussieht“ (28)
			Wiedererkennen Uhr	2*	

**Tabelle 2** Antworten zum Frageverständnis von 14 Patienten (Mehrfachantworten möglich; \* als „nicht verstanden“ gewertet)

die Alltagssicht in den Vordergrund stellen, geht es ihnen um Eigenverantwortung (z.B. zur Impfung: „gewissenhafter Umgang mit der eigenen Gesundheit“, P.38) und die Klärung ihrer Alltagskompetenzen bzw. des täglichen Hilfsbedarfs. Wenn die medizinische Sichtweise oben an steht, deuten Patienten die Fragen im Sinne einer diagnostischen Klärung, als Risikoabschätzung oder als Einschätzung des therapeutischen Bedarfs (z.B. zur Impfung: „fachgerechte Vorbeugung“, P.03). Dass Patienten ihre Alltagssituation bedenken, erweitert das Interpretationsspektrum der Fragen über die Ebene der rein medizinischen Sachbotschaft hinaus. Die Patienten interpretieren die Fragen im Sinne Schulz von Thuns [32] auch auf weiteren Ebenen. Zum Beispiel zeigen sich beim Frageverständnis zur Depression Selbstoffenbarung und Appell: „Leben alleine in der Wohnung“ (P.28) bzw. „kann Familie Rückhalt geben“ (P.20). Der Beziehungsaspekt wird bei einer Antwort zu den Hörfragen deutlich: „das Arztgespräch gut mitbekommen“ (P.35).

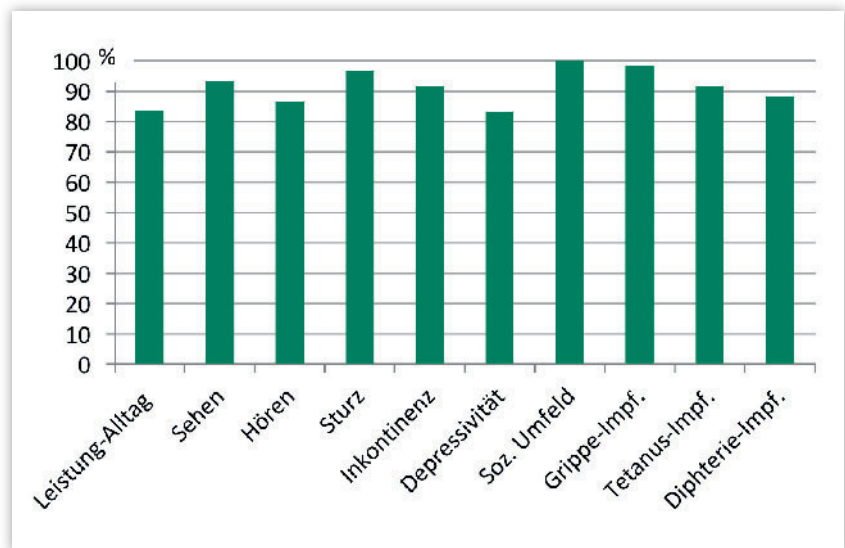
#### Änderungsvorschläge für das MAGIC-Assessment

Zwei Umformulierungen werden vorgeschlagen: 1.) Die Frage zur Leistungsfähigkeit bei Alltagsarbeiten ist kompliziert gestellt. Keiner der Patienten ist auf den erklärenden Teil der Frage „Hatten Sie aufgrund Ihres Gesundheitszustandes oder Ihrer Stimmung Schwierigkeiten ...“ im Probing eingegangen. Deshalb soll wegen der besseren Eingängigkeit dieser Zusatz weggelassen werden. 2.) Die erste Frage zum Sehen ist ebenfalls „holperig“ und wird deshalb umformuliert in: „Haben Sie Probleme, Zeitungsschrift zu erkennen – auch mit Brille?“

#### Stärken und Limitationen

Das Probing erweitert den bisherigen Erkenntnisstand zur Validität der Fragen hinsichtlich des Frageverständnisses der Patienten. Weiterhin ist das Instrument als Ganzes erstmalig auf die Reliabilität der Antworten in einem Test-Retest-Verfahren untersucht worden.

Unser methodisches Vorgehen weist jedoch auch einige Schwächen auf. Die Validität des MAGIC-Assessments ist durch unsere Untersuchung nicht umfänglich erfasst worden. Unterschiedli-



**Abbildung 1** Test-Retest-Reliabilität: Übereinstimmung der Testergebnisse (%) von 60 Patienten zu t0 und t1 (nach 1 Woche)

#### PD Dr. med. Ulrike Junius-Walker ...



... ist Fachärztin für Allgemeinmedizin und Forschungs Koordinatorin am Institut für Allgemeinmedizin der Medizinischen Hochschule Hannover.

che Gesichtspunkte, die Konstrukt-, Kriteriums- und Inhaltsvalidität, werden zur Erfassung der Testgüte benötigt. Wir haben uns nur auf das Frageverständnis, als einen wesentlichen Teil der Inhaltsvalidität bezogen [10]. Weiterhin hat sich gezeigt, dass das zugrunde liegende Konstrukt der Fragen vom professionellen Verständnis her auf die Erfassung von Körperfunktionen ausgerichtet ist, die Deutung aus Patientensicht aber gleichermaßen eine Alltagsbezogenheit enthält. Patienten haben die Fragen daher in ihrem Bedeutungsrahmen breiter ausgelegt. Das erschwert unsere Beurteilung darin, ob Patienten eine Frage „falsch“ verstehen – im Sinne von „außerhalb des Bedeutungsrahmens liegend“. Jedoch hat sich zugleich herausgestellt, dass beim MAGIC die subjektive alltagsbezogene Problembewertung der Patienten ergiebiger ist als die funktionell objektive. Wenn z.B. ein Patient unter „Leistungsfähigkeit bei Alltagsarbeiten“ angibt, er habe Schwierigkeiten im landwirtschaftlichen Betrieb mitzuarbeiten, dann geht es um seine persönliche ver-

hältnisbezogene Einschätzung im Kontext der spezifischen Lebensumstände.

Außerdem wurde die wiederholte Befragung nach einer Woche nicht unter absolut gleichen Voraussetzungen durchgeführt. Es wechselte sowohl der Ort der Befragung (erst Praxis, dann Wohnung des Patienten) als auch dessen Länge und die Reihenfolge der Items (erst Kurzform MAGIC, dann Langform STEP). Die Befragungen wurden von einem einzigen Interviewer vorgenommen. Im Praxisalltag mit verschiedenen Interviewern könnte die hier erzielte Übereinstimmung geringer ausfallen.

Letztlich müssen die Beurteilungen der Studienpatienten nicht unbedingt repräsentativ für alle älteren hausärztlichen Patienten sein, weil weder Ärzte bzw. Patienten randomisiert selektiert wurden, noch auf örtliche Gegebenheiten Rücksicht genommen wurde.

#### MAGIC: Anwendung in der Praxis

Mithilfe des MAGIC-Assessments können Hausärzte typische und relevante

Altersprobleme in systematischer Weise untersuchen. Diese Arbeit erweitert den Kenntnisstand zur Validität und Test-Retest-Reliabilität des MAGIC. Der hausärztliche Praxistest hat gezeigt, dass ältere Patienten die Fragen gemäß ihrer medizinischen Intention verstehen und zugleich auf den eigenen Lebenskontext beziehen. Der persönliche Alltagsbezug ist für Hausärzte besonders wichtig, weil sie Diagnostik- und Therapieentscheidungen im Kontext

patientenseitiger Erwartungen und Lebensumstände treffen. Die Testergebnisse des MAGIC sind weitgehend stabil und somit wenig abhängig von anderen Einflussfaktoren, wie z.B. Tagesform oder Untersuchungsort. Da ein Einsatz validierter Tests im geriatrischen Basisassessment gefordert wird, könnte das MAGIC-Assessment auch in diesem Rahmen eingesetzt werden. Es dient der Identifizierung geriatrisch vulnerabler Patienten und der Klärung

ihres persönlichen therapeutischen Bedarfes.

**Interessenkonflikte:** keine angegeben.

#### Korrespondenzadresse

PD Dr. med. Ulrike Junius-Walker  
Institut für Allgemeinmedizin  
Medizinische Hochschule Hannover  
Carl-Neuberg-Str. 1, 30625 Hannover  
Tel.: 0511 532-4928  
junius-walker.ulrike@mh-hannover.de

## Literatur

1. Böhm K, Tesch-Römer C, Ziese T. Gesundheit und Krankheit im Alter. In: Beiträge zur Gesundheitsberichterstattung des Bundes. Berlin: Robert Koch-Institut, 2009
2. Landefeld CS. Improving health care for older persons. *Ann Intern Med* 2003; 139: 421–424
3. Eichler K, Scrabal C, Steuerer J, Mann E. Preventive health risk appraisal for older people and impact on GPs' patient management: a prospective study. *Fam Pract* 2007; 24: 604–609
4. Thiem U, Hinrichs T, Müller CA, et al. Voraussetzungen für ein neues Versorgungsmodell für ältere Menschen mit Multimorbidität: Ergebnisse und Schlussfolgerungen aus 3-jähriger Forschung im „PRISCUS“-Verbund. *Z Gerontol Geriat* 2011; 44: 101–112
5. Piccoliori G, Gerolimon E, Abholz HH. Geriatric assessment in general practice using a screening instrument: is it worth the effort? Results of a South Tyrol Study. *Age Ageing* 2008; 37: 647–52
6. Sandholzer H, Hellenbrand W, Renteln-Kruse W, Van Weel C, Walker P. STEP-standardized assessment of elderly people in primary care. *Dtsch Med Wochenschr* 2004; 129: 183–226
7. Voigt I, Wrede J, Diederichs-Egidi H, Dierks M, Junius-Walker U. Priority setting in general practice: health priorities of older patients differ from treatment priorities of their physicians. *Croat Med J* 2010; 51: 483–492
8. Barkhausen T, Junius-Walker U, Hummers-Pradier E, Mueller CA, Theile G. "It's MAGIC" – development of a manageable geriatric assessment for general practice use. *BMC Fam Pract* 2015; 16: 4
9. Geyer S. Forschungsmethoden in den Gesundheitswissenschaften – Eine Einführung in die empirischen Grundlagen. Weinheim und München: Juventa Verlag, 2003
10. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity – establishing and reporting the evidence in newly-developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2 – Assessing Respondent Understanding. *Value Health* 2011; 14: 978–988
11. [www.kbv.de/media/sp/EBM\\_2015\\_01\\_01\\_BA\\_341\\_BeeG\\_EBM\\_Hauserztlich\\_geriatrisches\\_Basisassessment.pdf](http://www.kbv.de/media/sp/EBM_2015_01_01_BA_341_BeeG_EBM_Hauserztlich_geriatrisches_Basisassessment.pdf) (letzter Zugriff am 11.12.2015)
12. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *J Adv Nurs* 2003; 42: 57–63
13. Pohontsch N, Meyer T. Das kognitive Interview – ein Instrument zur Entwicklung und Validierung von Erhebungsinstrumenten. *Rehabilitation* 2015; 54: 53–9
14. Mayring P. Grundlagen und Techniken. In: Qualitative Inhaltsanalyse. Weinheim: Beltz Verlag, 10. Auflage, 2008
15. [www.ph3c.org/PH3C/docs/27/000150/0000103.pdf](http://www.ph3c.org/PH3C/docs/27/000150/0000103.pdf) (letzter Zugriff am 23.11.15)
16. McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires. Oxford: Oxford University Press, 2006: 89–93
17. Peach H, Green S, Locker D, et al. Evaluation of a postal screening questionnaire to identify the physically disabled. *Int Rehabil Med* 1980; 2: 189–193
18. Fuller G. Falls in the elderly. *Am Fam Physician* 2000; 61: 2159–68
19. Gassmann K, Ruppert R, Freiburger E. Predictors for occasional and recurrent falls in community-dwelling older people. *Z Gerontol Geriat* 2009; 42: 3–10
20. Anders J, Dapp U, Laub S, von Renteln-Kruse W, Juhl K. Einschätzung der Sturzgefährdung gebrechlicher, noch selbstständig lebender, älterer Menschen. *Z Gerontol Geriat* 2006; 39: 268–276
21. Sandvik H, Hunskaar S, Vanvik A, Bratt H, Seim A, Hermstad R. Diagnostic classification of female urinary incontinence: an epidemiological survey corrected for validity. *J Clin Epidemiol* 1995; 48: 339–343
22. Bosanquet K, Mitchell N, Gabe R, et al. Diagnostic accuracy of the Whooley depression tool in older adults in UK primary care. *J Aff Dis* 2015; 182: 39–43
23. Fillenbaum G, Smyer MA. The development, validity, and reliability of the Oars Multidimensional Functional Assessment Questionnaire. *J Gerontol* 1981; 36: 428–434
24. Burholt V, Windle G, Ferring D, et al. Reliability and validity of the Older Americans Resources and Services (OARS) Social Resources Scale in six European countries. *J Gerontol* 2007; 62: 371–379
25. George LK, Fillenbaum GG. OARS methodology: A decade of experience in geriatric assessment. *J Am Ger Soc* 1985; 33: 607–615
26. Laurence A, Lewis P, Gately C, Dixon A. Influenza and pneumococcal vaccination: do older people know if they have been vaccinated? *Aust NZ J Pub Health* 2016 (in press; Epub ahead of print: 10 Aug 2015)
27. Rolnick SJ, Parker ED, Nordin JD, et al. Self-report compared to electronic medical record across eight adult vaccines: do results vary by demographic factors? *Vaccine* 2013; 31: 3928–35
28. Watson YI, Arfken CL, Birge SJ. Clock completion: an objective screening test for dementia. *J Am Ger Soc* 1993; 11: 1235–40
29. Nishiwaki Y, Breeze E, Smeeth L, et al. Validity of the clock-drawing test as a screening tool for cognitive impairment in the elderly. *Am J Epidemiol* 2004; 160: 797–807
30. Turner S, Illife S, Downs M, et al. General practitioners' knowledge, confidence and attitudes in the diagnosis and management of dementia. *Age and Ageing* 2014; 33: 461–467
31. Pinto E, Peters R. Literature review of the clock drawing test as a tool for cognitive screening. *Dement Geriatr Cogn Disord* 2009; 27: 201–213
32. Schulz von Thun F. In: Miteinander reden – Störungen und Klärungen. Reinbeck bei Hamburg: Rowohlt, 2013