

AWMF-Register-Nr. 084-002LG

S1-Leitlinie
Geriatrisches Assessment der Stufe 2,
Living Guideline

Anmeldende Fachgesellschaft:
Deutsche Gesellschaft für Geriatrie



Anmelder bei der AWMF, Koordination

Erstanmeldung 2019: Sonja Krupp, Helmut Frohnhofen

Aktualisierungen 2021-2026: Sonja Krupp

Expertengruppe

Erstanmeldung 2019: Mitglieder (in alphabetischer Reihenfolge)

Jürgen Bauer, Michael Denking, Ellen Freiberger, Helmut Frohnhofen, Claudine Geser, Sabine Goisser, Werner Hofmann, Bernhard Iglseder, Michael Jamour, Sonja Krupp, Rainer Neubart, Caroline Renner, Ulrike Sommeregger, Stefan Strotzka, Walter Swoboda, Jens Trögner, Dorothee Volkert, Dirk K. Wolter

Aktualisierungen 2021-2026: AG Assessment der DGG mit Beiträgen von (in alphabetischer Reihenfolge)

2022: Michael Denking, Ellen Freiberger, Werner Hofmann, Bernhard Iglseder, Michael Jamour, Sonja Krupp, Caroline Renner, Sascha Sommer, Stefan Strotzka

2023: Michael Denking, Ellen Freiberger, Werner Hofmann, Sonja Krupp, Caroline Renner, Sascha Sommer, Stefan Strotzka, Dirk K. Wolter

2024: Petra Benzinger, Michael Denking, Ellen Freiberger, Werner Hofmann, Sonja Krupp, Caroline Renner, Sascha Sommer, Stefan Strotzka, Dirk K. Wolter

2025: Simone Brefka, Michael Denking, Ellen Freiberger, Michael Gagesch, Bernhard Iglseder, Sonja Krupp, Stefan Strotzka, Hans-Werner Wahl

2026: Petra Benzinger, Ellen Freiberger, Bernhard Iglseder, Sonja Krupp, Stefan Strotzka, Walter Swoboda

Beteiligte Fachgesellschaften

Die Deutsche Gesellschaft für Gerontologie und Geriatrie e. V. (DGGG), die Österreichische Gesellschaft für Geriatrie und Gerontologie (ÖGGG) und die Schweizerische Fachgesellschaft für Geriatrie (SFGG) haben sich mit Expertenbeiträgen an der Erarbeitung der Leitlinie beteiligt.

Die Vorstände der beteiligten Fachgesellschaften haben der Leitlinie in der vorliegenden Fassung zugestimmt.

Leitlinien sind für Ärzt*innen rechtlich nicht bindend [<https://www.awmf.org/regelwerk>, Zugriff 30.12.2025].

Danksagung

Die in der DGG verankerte AG Assessment mit Mitgliedern aus Deutschland, Österreich und der Schweiz bedankt sich für die Unterstützung bei der diesjährigen Aktualisierung der Leitlinie bei Prof. Dr. Erwin Scherfer und Dr. Johannes Trabert.

Methodik

S1 Leitlinie ohne systematische Literaturrecherche. Konsensusprozess im Umlaufverfahren. „Living Guideline“ mit jährlicher Aktualisierung. Aktueller Stand: 14.02.2026. Gültig bis 14.02.2027.

Die Interessenerklärungen aller an der Aktualisierung der Leitlinie Beteiligten wurden über das Portal „Interessenerklärungen online“ eingeholt. Die externe Bewerterin Frau Jennifer Kasper hat bei der Bewertung die von der AWMF vorgeschlagene Kategorisierung mit den entsprechenden Konsequenzen zugrunde gelegt:

Ausprägung Interessenkonflikt	Umstände für diese Kategorie	Konsequenz
Kein	-	-
gering	Einzelne Vorträge finanziert von der Industrie	Limitierung von Leitungsfunktion (Koordination/AG Leitung)
moderat	Tätigkeit in einem industriefinanzierten Advisory Board/Wiss. Beirat/als Gutachter Managementverantwortung industriefinanzierte Studie Aktienbesitz einzelner Firmen	Keine Abstimmung für die thematisch relevanten Empfehlungen oder Doppelabstimmung
hoch	Eigentumsinteresse Arbeitsverhältnis bei der Industrie Hoher Aktienbesitz einzelner Firmen	Keine Teilnahme an Beratungen und keine Abstimmung

Keine der Autor*innen hatte einen Interessenkonflikt mit Bezug zu den Inhalten der Leitlinie.

Für die vorliegende Leitlinie wurden im Auftrage und im Namen des Präsidiums der DGG (Deutsche Gesellschaft für Geriatrie e.V.) fachlich ausgewiesene Expert*innen in die Leitliniengruppe zur Er- oder Überarbeitung der S1 LL zum Geriatrischen Assessment der Stufe 2 berufen. Diese wurden - medizinisch wissenschaftlich selbstständig und unabhängig, jedoch in Ab- und Rücksprache mit der entsendenden Fachgesellschaft im Auftrag und im Namen dieser - unter Berücksichtigung des AWMF-Regelwerks tätig. Die vorliegende, von diesen Expert*innen entwickelte und im Konsens verabschiedete S1 Leitlinie ist urheberrechtlich als Gemeinschaftswerk anzusehen. Entsprechend der Zielsetzung und den Maßgaben der Leitlinienentwicklung der DGG und der AWMF soll die so entwickelte und verabschiedete Leitlinie über die AWMF und das Präsidium der DGG und deren Leitlinienbeauftragten in Fachkreisen und bei den Patient*innen Verbreitung finden. Zur Erreichung dieses Ziels wird die Leitlinie durch die Berechtigten (Autor*innen und der Vorstandsvorsitzende der AWMF zum Zeitpunkt des Inkrafttretens der Leitlinie) kostenlos abgegeben. Zur Unterstützung dieses Ziels müssen Dritte die Nutzungsrechte vertraglich mit den Berechtigten (AWMF und DGG) schriftlich gemäß Mustervertrag der AWMF vereinbaren, damit die Rechte dann unentgeltlich eingeräumt werden können, um deren gemeinnützige Zwecke zu fördern. Die Autor*innen als Urheber verstehen dies als ihren ehrenamtlichen Beitrag zur Verwirklichung der satzungsmäßigen Ziele der Berechtigten. Daneben ist es jedoch wegen des jederzeit möglichen Gewinns neuer medizinischer Erkenntnisse erforderlich, dass die AWMF und die DGG zu jeder Zeit die Kontrolle über die Verbreitung der Leitlinie behalten, um mitunter kurzfristig erforderliche Korrekturen oder auch Rückrufe der Leitlinie oder einzelner Bestandteile sicherzustellen. Unter keinen Umständen soll eine veraltete Leitlinie unter dem Namen der Berechtigten verbreitet werden.

Inhaltsverzeichnis

- 1 Einleitung 5
 - 1.1 Präambel 5
 - 1.2 Begriffsdefinitionen 5
 - 1.3 Begründung der Notwendigkeit der Leitlinie 8
 - 1.4 Ziele der Leitlinie..... 8
 - 1.5 Prinzipien der Instrumentenauswahl..... 9
 - 1.6 Patientenzielgruppe 12
 - 1.7 Versorgungsbereich 12
 - 1.8 Adressaten..... 12
- 2 Die Dimensionen 12
 - 2.1 Sehen / Hören..... 12
 - 2.2 Selbsthilfefähigkeit..... 14
 - 2.3 Mobilität / Motorik..... 20
 - 2.4 Kognition / Delirerfassung..... 46
 - 2.5 Depressive Störung 81
 - 2.6 Soziale Situation 93
 - 2.7 Schmerz..... 95
 - 2.8 Ernährung / Mundgesundheit / Dysphagie..... 100
 - 2.9 Kontinenz..... 109
 - 2.10 Schlaf..... 110
 - 2.11 Substanzmissbrauch/Sucht 111
- 3 Literatur..... 117
- 4 Abkürzungen..... 123

Bitte wie folgt zitieren:

Die Leitliniengruppe der AG Assessment der Deutschen Gesellschaft für Geriatrie e. V.,
S1-Leitlinie Geriatrisches Assessment der Stufe 2, Living Guideline,
Version 14.02.2026, AWMF-Register-Nr. 084-002LG

Zur Vereinfachung der Lesbarkeit nennen wir für Personen die weibliche Variante unter Verwendung des *. Sämtliche Personenbezeichnungen gelten jedoch unabhängig von der Geschlechtszugehörigkeit. Die Quellenangabe erfolgt im Text nur unter Nennung der Erstautor*in und des Jahres der Veröffentlichung, sofern dies zur Identifikation der Quelle ausreicht. Direkt auf eines der Assessment-Instrumente bezogene Quellennachweise finden sich in der jeweiligen Standardtabelle (Erstveröffentlichung unterstrichen), die weiteren in Kapitel 3.

1 Einleitung

1.1 Präambel

Geriatrische Patient*innen sind charakterisiert durch Einschränkungen der körperlichen und/oder geistigen Funktionalität bei Multimorbidität sowie durch ein erhöhtes Risiko für die Entwicklung von Gesundheitsstörungen und einen oft ungünstigeren, komplizierteren Verlauf. Monodimensionale und -kausale Ansätze können deshalb häufig keine optimale Besserung der Selbsthilfefähigkeit im Sinne von Aktivität und Partizipation gemäß der Internationalen Klassifikation der Funktionen der WHO (ICF) [WHO 2005] erreichen, auch wenn sie bei einem einzelnen Krankheitsbild evtl. erfolgreich sind. Die wesentlichen Faktoren zu identifizieren, die bei der Therapie berücksichtigt werden müssen, um die Patient*in in der Verbesserung und Erhaltung der Lebensqualität zu unterstützen, erfordert einen mehrdimensionalen Ansatz.

Dr. Marjory Warren erkannte in den dreißiger und vierziger Jahren des letzten Jahrhunderts zudem, dass Funktionsstörungen und Bettlägerigkeit nicht immer auf klassische Diagnosen zurückzuführen sind und entwickelte zum ersten Mal in der Medizingeschichte eine strukturierte geriatrische Untersuchung, welche Aspekte wie körperliche Funktionsfähigkeit, Inkontinenz und geistige Fähigkeiten beinhaltete [Warren 1943]. Das Geriatrische Assessment war geboren.

Im Laufe der Jahrzehnte hat sich das Geriatrische Assessment – konzeptionell auf der Idee Marjory Warrens basierend – inhaltlich, hinsichtlich der Gütekriterien und der Einsatzmöglichkeiten deutlich verändert und erweitert. Insbesondere nachdem Rubenstein und Kollegen 1984 zeigen konnten, dass sich die Qualität der ärztlichen Versorgung in den so genannten Veterans-Affairs Krankenhäusern durch Einführung eines Geriatrischen Assessments deutlich verbessern ließ, wurde das Assessment zu einem geriatrischen Standard. In der Folge entwickelte sich eine immer bessere Evidenzlage für unterschiedliche Dimensionen/Domänen [Rubenstein 1984]. Die vorhandenen Assessment-Instrumente wurden in Bezug auf ihre Gütekriterien untersucht und neue, geeignetere Verfahren eingeführt.

Die vorgelegte Leitlinie soll diese Entwicklung fortführen und eine Entscheidungshilfe für das geriatrische Basisassessment auf Grundlage der aktuellen Evidenz bieten. Der Fokus liegt dabei auf der Darstellung der einzelnen Instrumente. Dies ist der augenfälligste inhaltliche Unterschied zu der S3-Leitlinie „Umfassendes Geriatrisches Assessment bei hospitalisierten Patientinnen und Patienten“ (AWMF Register-Nr. 084-003). Die S1-Leitlinie „Geriatrisches Assessment der Stufe 2, Living Guideline“ bezieht darüber hinaus den ambulanten Sektor mit ein. Sie zielt auf Patient*innen, die bereits als „geriatrisch“ erkannt wurden. Dies grenzt sie von der S1-Leitlinie „Geriatrisches Assessment in der Hausarztpraxis“ (AWMF Register-Nr. 053-015) ab, die zudem ihren Schwerpunkt auf die Gestaltung mit den Besonderheiten von Hausarztpraxen kompatibler Abläufe legt.

1.2 Begriffsdefinitionen

Geriatrisches Assessment wird nach den Empfehlungen der AGAST [Arbeitsgruppe Geriatrisches Assessment 1995, Hofmann 1995] in drei Stufen eingeteilt:

Das Assessment der Stufe 1 entspricht einem multidimensionalen Screening mit dem Zweck der Identifikation geriatrischer Patient*innen. Dabei werden Instrumente

eingesetzt, die einen orientierenden Einblick in mehrere Problembereiche ermöglichen. Für dieses multidimensionale Screening wurde von der AGAST ausschließlich das Screening nach Lachs [Lachs 1990] in einer modifizierten Version genannt. Weitere Beispiele für Instrumente, die für die Stufe 1 eingesetzt werden, sind das Manageable Geriatric Assessment (MAGIC) [Junius-Walker 2016], die „Identification of Seniors at Risk“ (ISAR) [Thiem 2012], der Geriatrie-Check [Hobert 2018, Download Homepage BWKG], das von der Arbeitsgemeinschaft zur Förderung der Geriatrie in Bayern eingeführte „Geriatrisches Screening bei Klinikaufnahme“ (GSK) [Download Homepage AFGiB] und GeriNOT [Feist 2022].

Das Assessment der Stufe 2 entspricht dem „Basisassessment“, das in Deutschland in der akutgeriatrischen frührehabilitativen Komplexbehandlung gefordert wird (daraus ist jedoch keine Vorgabe ableitbar z. B. für Eintrittsassessments, die in der Schweiz im Rahmen der Akutgeriatrie oder der geriatrischen Rehabilitation verlangt werden).

Das Assessment der Stufe 3 wird bei Bedarf gezielt und „problemorientiert“ verwendet. Es benutzt „differenziertere und spezifische Verfahren“ zur Beantwortung spezieller Fragestellungen. Das bedeutet konkret, dass diese Verfahren immer einer klaren Indikation bedürfen und daher nicht bei allen Patient*innen angewendet werden sollen.

Die Begriffe „Screening“ und „Basisassessment“ reichen nicht zur klaren Abgrenzung der drei Stufen des Geriatrischen Assessments aus.

In Deutschland wird das Screening nach Lachs im Leistungskatalog der Krankenversicherungen über die Gebührenordnungsposition (GOP) 03360 im Einheitlichen Bewertungsmaßstab (EBM) dem „hausärztlich-geriatrischen Basisassessment“ zugeordnet. Diese Bezeichnung induziert Unklarheit. Die Leistung könnte ohne Stufenzuordnung verbleiben („Hausärztlich-geriatrisches Assessment“) oder dem multidimensionalen geriatrischen Screening zugeordnet und mit entsprechenden Inhalten gefüllt werden oder - unter Einbezug des hausärztlichen Vorwissens ggf. auf wenige Dimensionen beschränkt - die diagnostische Tiefe der Stufe 2 des geriatrischen Assessments angestrebt werden.

Andererseits dienen die Instrumente der Stufe 2 auch dem unidimensionalen Screening (= Untersuchung asymptomatischer Personen auf das Risiko, an der Zielstörung zu leiden [Spix 2012]), um noch nicht bekannte Störungen aufzudecken. Finden sich in Bezug auf eine bestimmte Dimension keine Beeinträchtigungen, so stellt dies eine Ressource dar.

Die Tabelle 1.2.1 definiert die Stufen des Assessments anhand der mit der Durchführung anvisierten Ziele. Beim Assessment der Stufe 2 lassen sich zwei Schritte unterscheiden: Der Ausschluss/Nachweis einer Störung (Stufe 2a) und die Erfassung der Ausprägung eines interventionsbedürftigen Problems (Stufe 2b). Einige schnell durchzuführende Instrumente sind zum Ausschluss einer therapielevanten Beeinträchtigung in der entsprechenden Dimension geeignet und erfüllen damit die Anforderungen an ein Assessment der Stufe 2a. Bei Hinweisen auf Therapiebedarf soll jedoch versucht werden, die Ausprägung der Beeinträchtigung genauer abzubilden und so die Stufe 2b des Assessments zu erreichen. Das Assessment der Stufe 2b kann aus mehreren Elementen bestehen, die sich in Kombination ergänzen. Auf die jeweilige Dimension bezogen, impliziert ein Assessment der Stufe 2b stets das der Stufe 2a und ein Assessment der Stufe 3 häufig das der Stufe 2b.

Die Tabelle 1.2.2 skizziert die drei Konstellationen, die sich nach dem Kennenlernen

dergeriatrischen Patient*in durch das interprofessionell agierende geriatriische Team aufgrund von anamnestischen Angaben, klinischen Hinweisen bzw. Untersuchungsbefunden und den Ergebnissen des Einsatzes von Assessmentinstrumenten im engeren Sinne (z. B. Fragebögen, Performance-Tests) ergeben können.

Tab. 1.2.1: Ziele des Geriatriischen Assessments nach Stufen

Stufe 1:	Identifikation geriatriischer Patient*innen
Stufe 2a:	Identifikation therapierelevant betroffener Dimensionen
Stufe 2b:	Dimensionsbezogene Beschreibung der Ausprägung von Beeinträchtigungen
Stufe 3:	Vertiefende Abklärung von Beeinträchtigungen

Tab. 1.2.2: Konstellationen im Geriatriischen Assessment der Stufen 2a, 2b und 3

Konstellation I: Kein Hinweis auf therapierelevante Beeinträchtigung in der Zieldimension gemäß Assessment der Stufe 2a UND Kein Hinweis auf therapierelevante Beeinträchtigung in der Zieldimension klinisch/anamnestisch gemäß Urteil des geriatriischen Teams → Assessment in der Zieldimension abgeschlossen
Konstellation II: Hinweis auf therapierelevante Beeinträchtigung in der Zieldimension gemäß Assessment der Stufe 2a ODER Hinweis auf therapierelevante Beeinträchtigung in der Zieldimension klinisch/anamnestisch gemäß Urteil des geriatriischen Teams → Durchführung des Assessments der Stufe 2b oder 3 in der Zieldimension
Konstellation III: Hinweis auf therapierelevante Beeinträchtigung in der Zieldimension gemäß Assessment der Stufen 2b oder 3 ODER Hinweis auf therapierelevante Beeinträchtigung in der Zieldimension klinisch/anamnestisch gemäß Urteil des geriatriischen Teams UND weiterer therapierelevanter Klärungsbedarf → gezielte Ergänzung durch weiteres Assessment

1.3 Begründung der Notwendigkeit der Leitlinie

Für alle Stufen des Geriatrischen Assessments sind neben den 10 von der AGAST genannten Instrumenten (Stufe 1: Geriatrisches Screening nach Lachs, Stufe 2: Barthel-Index, Mini Mental State Examination, Geriatrische Depressions-Skala, soziale Situation nach Nikolaus, Handkraft, Geldzählen nach Nikolaus, Timed Up and Go, Mobilitätstest nach Tinetti, Uhrentest nach Watson) inzwischen viele weitere verfügbar, die einen passgenaueren Einsatz ermöglichen oder zusätzliche Dimensionen in das Assessment einbeziehen. Eine Bewertung von deren Relevanz, Anwendbarkeit und Testökonomie sowie differenzierte Empfehlungen für den Einsatz dieser Instrumente fehlen bisher weitgehend.

Für Geriater*innen, die an geriatrischen Kliniken tätig sind, ist die Entwicklung einer Leitlinie für das geriatrische Assessment der Stufe 2 gegenüber den anderen Stufen vordringlich, da die Identifikation als geriatrischer Fall bereits durch die Bestätigung der Aufnahmeindikation abgeschlossen ist und das ggf. erforderliche Assessment der Stufe 3 auf dem der Stufe 2 aufbaut.

In der Kodierung der geriatrischen frührehabilitativen Komplexbehandlung [Bundesinstitut für Arzneimittel und Medizinprodukte 2024] ist das korrekte Durchführen des Assessments der Stufe 2 hinsichtlich der sozialen Situation, Selbsthilfefähigkeit, Mobilität, Kognition und Emotionslage für die Vergütung relevant. Assessment dient der Optimierung der Therapie. Ein Instrument, dessen Einsatz nicht weniger **therapierelevante** Information erbringt, als mit dem korrespondierenden von der AGAST genannten (Timed Up and Go, Barthel-Index, MMSE, GDS-15) zu erwarten gewesen wäre, ist **im konkreten Fall** (mindestens) gleichwertig.

1.4 Ziele der Leitlinie

Diese Leitlinie kann nur eine kleine Auswahl der im deutschsprachigen Raum bekannten Instrumente der Stufe 2 präsentieren, nicht genannte Assessmentverfahren sind nicht als unterlegen anzusehen. In Kapitel 2 werden die Instrumente dimensionsbezogen in einem kurzen Begleittext in Relation zueinander vorgestellt und in einer Tabelle in standardisierter Form charakterisiert. Beispielhaft werden Möglichkeiten aufgezeigt, den in der Therapie zunehmend angestrebten Ansatz der personalisierten Medizin auf das Assessment auszudehnen.

Das Assessment der Stufe 2 (a oder b) einschließlich der Berücksichtigung von Hinweisen zur Anamnese und klinischen Untersuchungsergebnissen aus dem Team benötigt einen mehrtägigen Zeitraum, soll jedoch sukzessive zeitnah komplettiert und bei Bedarf aktualisiert werden. Checklisten können die Erfassung aller vorgesehenen Dimensionen, Entscheidungsbäume den patientenzentrierten individuellen Einsatz der Verfahren erleichtern.

Bei der Zusammenstellung von Assessment-Instrumenten zu einem auf die Patient*in zugeschnittenen „comprehensive geriatric assessment“ muss es das Ziel sein, **im individuellen Fall** möglichst viel therapeutisch nutzbaren Erkenntnisgewinn auf möglichst ressourcenschonende, die Patient*in wenig belastende Weise zu erreichen. Hinweisen auf das Vorliegen einer Beeinträchtigung, die von (haus-)ärztlichen Kolleg*innen gegeben werden oder im Rahmen der vielfältigen Kontakte der Mitglieder des geriatrischen Teams mit den Patient*innen erkennbar werden, ist auch bei primär unauffälligen Testergebnissen nachzugehen. Liegt eine therapierelevante

Beeinträchtigung vor, so sollten Assessment-Instrumente eingesetzt werden, bei denen im Verlauf mit einer Änderung des Testergebnisses gerechnet wird, um einerseits Therapieerfolge, andererseits eine mögliche Progression der Störung abbilden zu können.

Geriat*innen in leitender Funktion soll die Leitlinie darin unterstützen, eine an der Zusammensetzung des Patientenkollektivs der Einrichtung orientierte Auswahl an Instrumenten zu treffen und Entscheidungsbäume für ihren standardisierten Einsatz festzulegen (Beispiele für eine patientenzentrierte Auswahl siehe Seiten 39, 69 und 88). Diese dimensionsbezogenen Algorithmen können als Basis der innerbetrieblichen Schulung des interprofessionell das Assessment durchführenden geriatrischen Teams dienen und zugleich in der Kommunikation mit dem Medizinischen Dienst und Kostenträgern für Transparenz sorgen.

Die heterogenen Aus- und Fortbildungsgänge der im Geriatrie-Team tätigen Professionen führen dazu, dass sich die vorliegende Qualifikation hinsichtlich eines sicheren Umgangs mit bestimmten Assessment-Instrumenten nicht an der Bezeichnung des Berufs- bzw. Studienabschlusses allein ablesen lässt. Die Leitlinie nimmt darauf durch die Einführung des Begriffs „spezifisch geschultes Fachpersonal“ Rücksicht. Es obliegt der für das Team verantwortlichen Geriat*in, ggf. den Bedarf für weitere Schulung zu erkennen und diese zu veranlassen oder durchzuführen.

1.5 Prinzipien der Instrumentenauswahl

Für die Durchführung des Assessments der Stufe 2 in den Bereichen Selbsthilfefähigkeit, Mobilität, Kognition und Emotion unter Anwendung des Barthel-Index, des Timed Up and Go, der MMSE und der GDS-15 wurde von der AGAST ein Zeitbedarf von 30-35 Minuten veranschlagt. Hinzu kommt im Rahmen frührehabitativer Komplexbehandlung obligat ein soziales Assessment. Wegen ihrer Bedeutung für die Lebensqualität und Prognose geriatrischer Patient*innen empfiehlt die Leitliniengruppe auch eine generelle Erfassung der Dimensionen Ernährung (einschließlich Dysphagie) und Schmerz in (ggf. hausintern) standardisierter Form sowie weiterer Dimensionen.

Die Auswahl der in dieser Leitlinie genannten Instrumente erfolgte geleitet von Gütekriterien. Diese werden daher, wo immer möglich, in den entsprechenden Tabellen aufgeführt. Die Kategorisierung eines Instruments als „geeignet“ für einen bestimmten Schweregrad wurde unter zwei Gesichtspunkten betrachtet: Die Eignung sollte sowohl hinsichtlich der Durchführbarkeit (feasability) als auch hinsichtlich der Erfassung des angegebenen Schweregrades gegeben sein. Die Leitlinie strebt an, auf alle im deutschen Sprachraum breit eingesetzten Instrumente für das Assessment der Stufe 2 einzugehen, beschränkt sich jedoch nicht auf diese [Kudelka 2024] - denn das würde ungewollt Stagnation fördern im angestrebten fortlaufenden Verbesserungsprozess.

Im Update 2026 wurden neben Literaturhinweisen insbesondere die Angaben hinsichtlich des Zeitbedarfs für die Durchführung der vorgestellten Instrumente überprüft und ggf. angepasst. Die Angaben beziehen sich auf den Zeitbedarf der untersuchten Personen einerseits, erfahrener Untersuchender andererseits und beziehen Wege zum Untersuchungszimmer nicht ein. Der reale Zeitbedarf wird unter anderem durch das allgemeine Kommunikationsverhalten der untersuchten Person wesentlich beeinflusst.

Zur Erhöhung der Transparenz bezüglich der Auswahl der in dieser Leitlinie mittels einer

Standardtabelle charakterisierten Instrumente wurde im Update 2024 der Absatz **„Begründung der Aufnahme in die Leitlinie“** eingeführt. Ausgehend von der 1995 von der AGAST vorgeschlagenen Testbatterie werden zur Abstimmung über die Aufnahme anstehende Instrumente hinsichtlich ihrer Eignung für geriatrische Patient*innen unter folgenden Aspekten betrachtet:

- Füllt das Instrument eine noch nicht durch die enthaltenen Instrumente abgedeckte diagnostische Lücke? Ist es geeignet, zusätzliche therapierelevante Informationen zu generieren?
- Zeigt es unter bestimmten Voraussetzungen (Setting, individuelle Patienteneigenschaften, z. B. Komorbiditäten) ein günstigeres Aufwand-Nutzen-Verhältnis als die bereits enthaltenen Instrumente?
- Ist es unter bestimmten Voraussetzungen hinsichtlich anderer Gütekriterien (z. B. Boden-/Deckeneffekte, diagnostische Genauigkeit, Reliabilität, Änderungssensitivität) überlegen?

Die Leitlinien-Gruppe verzichtet bewusst auf die Darstellung von Instrumenten, die mit einer fixen Kombination von Items mehrere Dimensionen abbilden sollen, da dies einer individuellen Anpassung an die Heterogenität geriatrischer Patient*innen nicht gerecht wird. Die in der Leitlinie dargestellten Assessment-Instrumente sollen stattdessen flexibel kombiniert werden, um den Mitgliedern des multiprofessionellen geriatrischen Teams mittels der erhobenen Befunde eine individualisierte Therapieplanung zu ermöglichen.

Hier soll auch konkret auf die weiteren Argumente für die Entscheidung gegen die Aufnahme des „Selbstständigkeits-Index für die Neurologische und Geriatrische Rehabilitation“ (SINGER) [Gerdes 2012] eingegangen werden, der im Rahmen des Intensivpflege- und Rehabilitationsstärkungsgesetzes (IPREG) und der damit geänderten Rehabilitations-Richtlinie als Ersatz für eine Überprüfung durch die Krankenkasse angewendet werden kann:

1. Das Instrument wurde nur im Setting der neurologischen Rehabilitation ausreichend hinsichtlich seiner Gütekriterien validiert, nicht prospektiv in einer für die spezifisch geriatrische Versorgung typischen Population. In dieser überwiegen nicht-neurologische Hauptdiagnosen sehr deutlich, so dass eine Übertragbarkeit nicht gegeben ist. Die Dominanz neurologisch-rehabitativer Aspekte bei der Entwicklung zeigt sich schon in der Verbalisierung der Stufenzuordnung (z. B. „Rückkehr“ von Erinnerungen, Item XVI).
2. Es verbindet die Anteile zu verschiedenen Dimensionen/Domänen (Selbstversorgung und Mobilität, Kommunikation, kognitive Aktivität) in einer fixierten Kopplung miteinander, diese Leitlinie strebt dagegen für jede einzelne Dimension eine an Patienteneigenschaften und Therapieziele adaptierte Instrumentenauswahl an.
3. In der Dimension Selbsthilfefähigkeit (SINGER: Selbstversorgung und Mobilität) wird die diagnostische Tiefe der Stufe 2b angestrebt, in der Dimension Kognition jedoch nicht erreicht. Dadurch stellt die Erhebung der Items zur Kognition einen zusätzlichen Aufwand dar, da außerdem (mindestens) ein weiteres Instrument für das Assessment der Kognition eingesetzt werden müsste. Die „Belegaufgaben“ sind wiederum nicht validiert.
4. Die Dimension Selbsthilfefähigkeit in Basisaktivitäten wird bereits durch ein besser validiertes Instrument, den Barthel-Index, abgedeckt. Die immer vorhandenen Nachteile der Einführung weiterer Instrumente (u. a. Lernaufwand, Verminderung der

Vergleichbarkeit mit Ergebnissen breiter etablierter Verfahren im individuellen Verlauf oder zu wissenschaftlichen Zwecken) werden nicht aufgewogen.

5. Der zeitliche und personelle Aufwand für die Erhebung ist erheblich höher.

6. Für den Einsatz des SINGER ohne logopädische und neuropsychologische Mitwirkung, im vertragsärztlichen oder Akutbereich liegen keine Gütekriterien vor. Der Einsatz des Instruments zur Indikationsstellung für geriatrische Rehabilitation ist derzeit nicht wissenschaftlich begründbar.

Der besser validierte Functional Independence Measure (FIM) [Keith 1987] wurde trotz der Möglichkeit, die Skalen für Selbsthilfefähigkeit (13 Items) und Kognition (5 Items) getrennt zu nutzen, ebenfalls zugunsten des breit eingeführten Barthel-Index nicht mit Standardtabelle in diese Leitlinie aufgenommen.

Diese Leitlinie beschränkt sich weitgehend auf die Nennung von Verfahren, mit denen in der Gruppe der Autor*innen auch praktische Erfahrungen vorliegen, da die Anwendbarkeit unter den realen Bedingungen der geriatrischen Versorgung in die Beschreibung einfließen sollte. Die Instrumente werden in der Regel nach Zeitbedarf in aufsteigender Reihenfolge präsentiert, an einigen Stellen wird zugunsten einer logischeren Reihenfolge davon abgewichen. Der angegebene mittlere Zeitaufwand für die Patient*in einerseits, die Untersucher*in andererseits (inkl. Vor- und Nachbereitung) ist als Anhaltspunkt zu verstehen und basiert sowohl auf Literaturangaben als auch auf persönlicher Erfahrung; die Spannweite ist groß.

Die Reihenfolge der Nennung der Ermittlung von Gütekriterien dienenden Studien erfolgt nach Aktualität. Bei der Angabe von Testresultaten in Punkten wird auf das Wort „Punkte“ verzichtet, Einheiten werden nach dem Système international d'unités (SI) abgekürzt.

Sofern ein Cut-off angegeben ist, entspricht die Zahl vor dem Schrägstrich dem letzten normalen Wert, die Zahl hinter dem Schrägstrich dem ersten auf eine Störung hinweisenden Wert.

Der Grundgedanke der Evidenzbasierung sollte auch für die Anwendung von Instrumenten gelten, die hier nicht enthalten sind. In die Entscheidung für ein Assessment-Instrument bei einer bestimmten Patient*in ist jedoch auch einzubeziehen, dass die geriatrietypische Multimorbidität sich nicht zuletzt im Vorliegen von Merkmalen wie Seh-, Hör- oder feinmotorischen Störungen manifestieren kann, die die Aussagefähigkeit vieler Testverfahren einschränken können. In der Beschreibung der in dieser Leitlinie genannten Instrumente werden Hinweise dazu im Absatz „Limitationen“ genannt. Seh- und Hörfähigkeit sind a. eigenständige Dimensionen hoher Relevanz für die Lebensqualität, b. prognostische Faktoren für die Entwicklung geriatrischer Syndrome und c. müssen bei der Auswahl von Assessment-Instrumenten auf der individuellen Ebene sowie d. in der Therapie Berücksichtigung finden.

Generell ist von einer reduzierten Belastbarkeit und Konzentrationsdauer geriatrischer Patient*innen auszugehen, so dass bei Nicht-Unterlegenheit hinsichtlich der therapierelevanten Aussagefähigkeit im konkreten Fall das schneller durchzuführende Instrument vorzuziehen ist – zumal der Störfaktor „Erschöpfung“ dabei weniger zu Buche schlägt und darüber hinaus Personalressourcen geschont werden.

Insbesondere das Assessment der Stimmung, in zweiter Linie das der Kognition liefert häufig nicht repräsentative Ergebnisse, wenn es bereits am ersten oder zweiten Behandlungstag vorgenommen wird, da ein Vertrauensverhältnis zwischen Patient*in und Untersucher*in Voraussetzung für eine aussagefähige Erhebung ist [Lezak 2012].

1.6 Patientenzielgruppe

Geriatrische Patient*innen.

1.7 Versorgungsbereich

Geriatrische Versorgung stationär, teilstationär, ambulant und mobil.

1.8 Adressaten

In spezifisch geriatrischen Versorgungsstrukturen tätige, für das Assessment verantwortliche Ärzt*innen und andere Mitglieder des geriatrischen Teams.

2 Die Dimensionen

Teil der jährlichen Aktualisierung dieser Leitlinie ist eine Überprüfung der behandelten Dimensionen (= Domänen) auf Vollständigkeit und bedarfsgerechte Berücksichtigung.

Assessment wird in hohem Maße durch die Kommunikationsfähigkeiten von (Untersuchenden und) Untersuchten beeinflusst. Hier fließen neben Sehen und Hören auch z. B. Werkzeugstörungen unterschiedlicher Genese sowie Defizite in der motorischen Steuerung der an der Kommunikation beteiligten Organe ein. Perspektivisch ist die Entwicklung eines Instruments wünschenswert, das, ausgehend von der Kommunikationsfähigkeit, innerhalb weniger Minuten eine grobe Differenzierung hinsichtlich der Genese und des Schweregrades der Beeinträchtigung in den wichtigsten Subdimensionen der Informationsaufnahme und -übermittlung – einschließlich des Sehens und Hörens – ermöglicht.

Ein geriatrisches Assessment wäre ohne die ausführliche klinische und medikamentöse Anamnese inkomplett, jedoch empfiehlt diese Leitlinie dafür kein bestimmtes Erhebungsinstrument. Mit der Erfassung und Beurteilung der Medikation befasst sich u. a. die Hausärztliche Leitlinie Multimedikation [AWMF Register-Nr. 053-043].

Auf die Vorstellung von Instrumenten zur Erfassung von Frailty wird trotz reichlichen Angebots in dieser Leitlinie bewusst verzichtet. Zwar ist den stark divergenten Definitionen gemeinsam, eine Personengruppe mit erhöhtem Risiko für negative Outcomes zu markieren, jedoch setzen Diagnostik und Therapie an den verschiedenen Dimensionen an. Die Leitlinie befasst sich daher ausschließlich direkt mit diesen.

2.1 Sehen / Hören

Beeinträchtigungen des Sehens und Hörens gehören nicht nur zu den geriatritypischen Syndromen, deren Aufdecken ggf. zu weiterer Diagnostik und Therapie führen soll, sondern ihr Vorliegen muss auch bei der Auswahl anderen Dimensionen zuzuordnender Assessment-Instrumente berücksichtigt werden, wenn die Ergebnisse valide sein sollen. Sensorische Defizite können u. a. zu einem geringeren kognitiven Input, einer Beeinträchtigung der (insbesondere instrumentellen) Aktivitäten des täglichen Lebens, der Kommunikation, der Stimmung und zu einem Anstieg des Sturzrisikos führen. Bei bereits vorhandenen kognitiven Defiziten ist die Optimierung der Sensorik von

besonderer Relevanz.

Laut einer britischen Studie [Bowen 2016] fanden sich bei einem Drittel der untersuchten Menschen mit Demenz relevante Visusverluste (ggf. trotz Verwendung der vorhandenen Brille), wobei Bewohnende in Pflegeeinrichtungen doppelt so häufig betroffen waren wie nicht institutionalisiert wohnende Personen. Der Hälfte der sehgestörten Patient*innen konnte durch eine Brillenversorgung ausreichend geholfen werden, bei einem weiteren Viertel war eine Katarakt-Operation indiziert. Das Aufdecken unkorrigierter Visusstörungen als Teil des physischen Assessments kann somit oft die Lebensqualität geriatrischer Patienten verbessern.

Auch außerhalb spezifisch geriatrischer Versorgung gehört eine orientierende Beurteilung des Gesichtsfeldes, der Bulbus- und Pupillomotorik sowie die Frage nach z. B. trockenen, tränenden oder schmerzenden Augen zur umfassenden ärztlichen Aufnahmeuntersuchung. Im Assessment der Stufe 2 gehen Informationen zu den verschiedenen Dimensionen in der Regel über die in der Stufe 1 - des multidimensionalen Screenings zur Detektion geriatrischer Patient*innen - hinaus. Ein standardisiertes Prozedere hierzu ist jedoch bislang nicht etabliert.

Die deutsche Übersetzung der ISAR [Thiem 2012] enthält die Frage „Haben Sie unter normalen Umständen erhebliche Probleme mit dem Sehen, die nicht mit einer Brille korrigiert werden können?“. Im Sinne einer für möglichst viele Patienten verständlichen Ausdrucksweise ist die vereinfachte Frage „Haben Sie Probleme mit dem Sehen?“ als Einstieg gut zu nutzen, die Brillenversorgung und -verwendung ggf. anschließend zu erfragen. Kognitiv unbeeinträchtigten Patient*innen, die keine Probleme mit dem Sehen angeben, sind Selbsttests im Intervall zwischen den allen Älteren anzurathenden jährlichen ophthalmologischen Kontrollen zu empfehlen. Hierfür stehen z. B. der ZEISS Online-Seh-Check (Landolt-Ringe für Sehschärfe und Kontrastsehen, Ishihara-Farbsehtest, Astigmatismus-Check, Amsler Gitter) und ACTO-Sehtest (Lesetest + Amsler Gitter) zur Verfügung.

Gibt ein*e Patientin Probleme mit den Augen an, so gilt es, den Stand des Einbezugs ophthalmologischer Kolleg*innen festzustellen bzw. diesen herzustellen („Waren Sie damit bereits bei eine*r Augenärzt*in?“ „Wann zuletzt?“ „Sollen Sie wiederkommen?“) Hat sich das Sehen nach der letzten Augen-Untersuchung verschlechtert, ist umgehend die Augenärzt*in zu informieren.

Zum Screening nach Lachs gehörte im Original [Lachs 1990], eine seitengetrennte Überprüfung des Nahvisus mit einer Sehtafel, bei Bedarf mit Brille. In der deutschen Übersetzung wurden daraus neben der Frage „Hat sich Ihre Sehfähigkeit in letzter Zeit verschlechtert?“ die beiden Aufgaben (1.) Fingerzahl aus 2 m Entfernung erkennen und (2.) Nahvisus/Lesen einer Überschrift. Mindestens bei subjektiv vorhandenen, nicht durch eine Brille kompensierten Sehstörungen außerhalb augenärztlicher Kontrolle sowie bei kognitiv gestörten, jedoch noch testbaren Patient*innen sollte eine orientierende Untersuchung erfolgen. Sofern keine Anhaltspunkte für ein anderes Problemfeld vorliegen, ist die Einschätzung des Nahvisus in der Regel ausreichend. Der Sehtest sollte seitengetrennt (Abdecken eines Auges) erfolgen im subjektiv am besten geeigneten Abstand. Weicht dieser erheblich von der üblichen Distanz (etwa die Länge eines DIN A4-Blattes) ab, so kann dies ein Hinweis auf nicht ausreichend kompensierte Refraktionsanomalie sein.

Sofern im Verlauf zunehmend schwerer Demenz (oder z. B. bei Fremdsprachigkeit,

Analphabetismus oder Alexie) Texte nicht gelesen werden können, erscheint es logisch, als Symbole im Sehtest Ziffern zu verwenden, da sowohl das Erkennen einzelner Buchstaben als auch die Beschreibung der räumlichen Ausrichtung von Zeichen (Landolt-Ringe, Snellen-E) kognitiv meist anspruchsvoller sein dürften. Es besteht im Hinblick auf die Verwendung von Sehtafeln bei kognitiv eingeschränkten Patient*innen jedoch noch Forschungsbedarf.

Für das Hören sieht das Screening nach Lachs einen seitengetrennten Test vor (ein Ohr zuhalten), bei dem aus etwa 50 cm Entfernung drei Zahlen nach Ausatmung geflüstert werden, von denen mindestens 2 erkannt werden sollen. Dieser Test erscheint zusätzlich zur Frage nach subjektiven Hörstörungen und der Beobachtung der Kommunikationsfähigkeit während der Anamnese und klinischen Untersuchung sinnvoll.

Der **Mini-Audio-Test (MAT)**, ein sechs Fragen enthaltender Fragebogen mit dreistufiger Selbsteinschätzung, ist niedrigschwellig einsetzbar. Bei Personen zwischen 60 und 75 Jahren erreicht er eine Spezifität von 80,1%, jedoch mit einer Sensitivität von 47,0% einen deutlich niedrigeren Wert als bei Jüngeren [Löhler 2019].

Das ebenfalls dreistufig skalierte **Hearing Handicap Inventory for the Elderly (HHIE)** [Ventry & Weinstein 1982, deutsche Übersetzung Bertoli 1996] mit 25 und seine Kurzform (HHIE-S) mit 10 Fragen bilden die emotionalen und sozialen Auswirkungen von Hörverlusten ab. Die Erfassung dieser Zusammenhänge kann - über den Hinweis auf den Bedarf an weiterer Diagnostik hinaus - für die Förderung der Motivation von betroffenen Patient*innen genutzt werden, ihre Hörprobleme anzugehen [Moser 2017].

Als Selbsttest stehen telefonisch oder online durchführbare Hörtests zur Verfügung, die teilweise auch ohne Kopfhörer durchführbar sind.

Liegen schwerwiegende Einschränkungen des Sehens oder Hörens vor, so ist dies bei der Auswahl der weiteren Assessment-Instrumente zu berücksichtigen.

2.2 Selbsthilfefähigkeit

Im Basisassessment der Selbsthilfefähigkeit ist stets der Einsatz eines Instrumentes der Stufe 2b erforderlich, da geriatrische Patient*innen generell durch zumindest drohende Einschränkungen ihrer selbständigen Lebensführung gekennzeichnet sind.

Die Leitliniengruppe empfiehlt derzeit kein alternatives Instrument zu dem von der AGAST vorgeschlagenen **Barthel-Index (BI)** nach Barthel und Mahoney (Tabelle 2.2.1) als überlegenen Ersatz. Auch im ambulanten Bereich gehört der Barthel-Index zu den vier Instrumenten, deren Eignung für die Erfassung der Selbstständigkeit in den täglichen Basisaktivitäten am besten belegt ist [Hopman-Rock 2018].

Das **Hamburger Manual zum Barthel-Index** [Lübke 2001] hat einen Teil der im Original gegebenen Definitionslücken geschlossen. Die Länge des zu beurteilenden Zeitraumes wurde von Barthel und Mahoney nicht festgelegt. Das Hamburger Manual formuliert in Bezug auf sämtliche Items eine mehrtägige Spanne, ohne die genaue Länge festzulegen („Sollten (z.B. je nach Tagesform) stets unterschiedliche Einstufungskriterien zutreffen, ist die niedrigere Einstufung zu wählen.“), für die Stuhlkontinenz sogar einen mehrwöchigen Zeitraum (Item 9: „durchschnittlich mehr als 1x/Woche“). Der Vergleich der Einstufungsergebnisse bei jeweils über 2600 Patient*innen geriatrischer Krankenhäuser vor und anderer Patient*innen der gleichen Krankenhäuser nach Einführung des Hamburger Manuals wurde mit der Langversion durchgeführt und ergab

eine hohe Korrelation der Summenscores sowohl bei Aufnahme als auch Entlassung trotz z. T. signifikanter Abweichungen bei Vergleich auf Itemebene [Lübke 2004]. Die meisten der diversen Formulierungen der Schweregrade von Beeinträchtigungen in den 10 erfassten Basisaktivitäten orientieren sich im deutschsprachigen Raum jedoch an der Kurzversion des Hamburger Manuals. Der **Barthel plus** [Krupp 2025] basiert auf der Langversion des Hamburger Manuals, mit dessen Summenscore er eng korreliert, schließt weitere Definitionslücken, ist sprachlich vereinfacht und erfasst auf Itemebene bei voller Punktzahl, ob trotz selbstständiger Durchführung eine Beeinträchtigung vorliegt („b“) oder nicht („a“). Für die Untersuchung der Gütekriterien wurde der Beurteilungszeitraum auf 72 Stunden (für die Stuhlkontinenz: 4 Wochen) festgelegt.

Der **Frühreha-Barthel-Index** [Schönle 1995] ergibt sich rechnerisch aus dem Barthel-Index und berücksichtigt bis zu sieben Situationen mit erhöhter Überwachungspflicht in Form von Minuspunkten (z.B. beaufsichtigungspflichtige Schluckstörung → -50 Punkte).

Der **Functional Independence Measure (FIM)** [Keith 1987] bietet durch seine siebenstufige Skalierung mehr Raum für Änderungssensitivität als der BI, lässt sich aber nicht zum Vergleich mit dem weitaus breiter etablierten Barthel-Index zur Verlaufsbeobachtung in diesen umrechnen.

Die **IADL-Skala** (Instrumental Activities of Daily Living) (Tabelle 2.2.2) schätzt die persönliche Kompetenz in der Lebensführung ein und umfasst 8 Items: Telefonbenutzung, Einkaufen, Essenzubereitung, Haushaltsführung, Wäsche, Benutzung von Transportmitteln, Einnahme von Medikamenten und Management der eigenen Finanzen. Die Informationen werden durch die Befragung der Betroffenen, gegebenenfalls auch der Bezugspersonen, gewonnen. Die Fragen sind gut verständlich, der Test leicht durchführbar. Die IADL-Skala ist sowohl im stationären (bezüglich des Status vor Aufnahme) als auch im ambulanten Bereich einsetzbar, eher nicht in Pflegeeinrichtungen. Die angesprochenen Bereiche sind entscheidend für eine selbstständige Lebensführung. Insofern kann die IADL-Skala als Erweiterung der die Basisaktivitäten erfassenden Selbsthilfe-Instrumente angesehen werden. In mehreren angelsächsischen Publikationen wird der Test als Checkliste empfohlen.

Die Leistung im **Geldzähltest nach Nikolaus** (Tabelle 2.2.3) als Performance-Test einer instrumentellen Aktivität hängt sowohl von kognitiven als auch feinmotorischen und sensorischen Fähigkeiten (Sehen, Tastsinn) ab. 9,80 Euro (ursprünglich 19,80 Deutsche Mark) befinden sich in standardisierter Stückelung in einem halbstandardisierten Portemonnaie, aus dem sie möglichst schnell entnommen und gezählt werden sollen. Sowohl bei dreimaliger Angabe einer falschen Summe als auch, wenn kein Resultat innerhalb von 5 Minuten genannt wird, sollen gemäß Originalarbeit 300 Sekunden als Testergebnis notiert werden. Ein Normalbefund (unter 45 Sekunden) spricht dafür, dass diese Person einfache Geldgeschäfte tätigen kann und wahrscheinlich auch ähnlich anspruchsvolle Tätigkeiten beherrscht, z. B. die korrekte Einnahme von Medikamenten nach Plan.

Tab. 2.2.1: Dimension Selbsthilfefähigkeit

Name des Instrumentes: Kürzel:	Barthel-Index BI	
Mittlerer Zeitaufwand: Für Patient*in 8 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 10 min		
Lernaufwand für Untersucher*in:		
O leicht	x mittel	O anspruchsvoll
Durchführung durch: Pflegekraft, Ärzt*in, Ergotherapeut*in. Während (teil-) stationärer geriatrischer Therapie interprofessionelle Erhebung unter Mitwirkung von Physiotherapeut*in sinnvoll.		
Geeignet für Schweregrad:		
O leicht/Früherkennung	x mittel	x schwer
Gütekriterien: Krupp (2025): n = 124, stationäre Akutgeriatrie, 81 ± 7 (60-101) J, 67% F. Barthel-Index zu t ₀ 62 ± 17 (10-100), kein signifikanter Unterschied der Summenscores bei Operationalisierung gemäß Hamburger Manual versus Barthel plus. Barthel plus Interrater-Reliabilität 0,97, interne Konsistenz Cronbachs Alpha 0,98, Änderungssensitivität über Cohens d 0,92 bei 13,6 ± 2,5 Tagen Intervall. Bei Items mit vollem Punktwert, aber Zusatz „b“ (beeinträchtigt) erhöhtes Risiko für Funktionsverlust (z. B. relatives Risiko 2,77 für Punktabfall innerhalb von 2 Monaten nach Krankenhaus-Entlassung). Bouwstra (2018): Verwendung einer niederländischen Übersetzung mit maximal 20 Punkten. Stationäre geriatrische Rehabilitation in spezialisierten Pflegeheimen, 2 Stichproben: 1. n = 207, 80 ± 8 J, 58% F, im Median 35 (22-55) Tage Verweildauer, interne Konsistenz (n = 199) Cronbachs Alpha 0,83, mittlere Änderung des modifizierten Summenscores von 4,8 ± 4,0 von 20 Punkten, Minimal Important Change auf Basis des Entlassortes (n = 177) 3,1 von 20 Punkten; 2. n = 37, 74 ± 11 J, 61% F, Interrater-Reliabilität 0,96, berechnet auf Basis des Urteils von 3 Bewerter-Paaren, die jeweils 7 bis 20 Bewertungen vornahmen. Hartigan (2011): Studienpopulation n = 65, 81,5 ± 7,5 (65 – 100) J, 62% F, konsekutiv aufgenommene Patient*innen unter Ausschluss höhergradig kognitiv beeinträchtigter Personen, stationäre Rehabilitation. Innerhalb von 5 Tagen je zweimal Bestimmung eines modifizierten BI (0 – 20 Punkte) durch Ärzt*innen sowie Pflegefachpersonen (PFP), die die Aktion beobachten konnten. Interrater-Reliabilität (Summenscore) Ärzt*in/PFP 0,6968 (Spearman), Ärzt*in/Ärzt*in 0,74, PFP/PFP 0,8290, wobei Ärzt*innen häufig die Fähigkeiten überschätzten. Niedrigste Interrater-Reliabilität im mittleren Bereich (10-16/20). Übereinstimmung auf Item-Ebene >0,61 Ärzt*in/PFP bei 3 Items, Ärzt*in/Ärzt*in bei 5 Item, PFP/PFP bei allen Items. Sainsbury (2005): Laut Review nur 3 von 12 Studien zur Reliabilität, die Personen ab 65 Jahren einschlossen, mit zufriedenstellender statistischer Analyse. Interrater-Reliabilität (ICC) 0,95-0,97. 95%-Konfidenz-Intervall ±4 Punkte. Korner-Bitensky (1995): n = 366, 25 - 98 J, ambulante Patient*innen nach Entlassung aus orthopädischer oder neurologischer Rehabilitation. Bei Vergleich von persönlichem mit		

telefonischem Interview Reliabilität (ICC) 0,89.

Begründung der Aufnahme in die Leitlinie:

Von der AGAST empfohlenes, international breit eingesetztes Instrument, dessen Gütekriterien bezüglich der „basic activities of daily living“ bislang von keinem anderen im geriatrischen Setting nachweislich übertroffen werden.

Limitationen / Forschungsbedarf:

Beeinträchtigungen, die kompensiert werden können (z. B. durch Hilfsmiteileinsatz) werden nicht im Score erfasst. Die Selbsthilfefähigkeit im Bereich instrumenteller Aktivitäten (z. B. Haushaltstätigkeiten) wird nicht abgebildet. Je kürzer der Beobachtungszeitraum, desto größer die wahren Schwankungen des der Beurteilung zugrundeliegenden Status, da sich dieser sich am Leistungstief innerhalb dieses Zeitraums orientiert (und daher desto niedriger die Retest-Reliabilität). Es besteht der Bedarf, einen Beobachtungszeitraum festzulegen, auf den sich die Angaben beziehen.

Literatur:

Deutsche Version unter
<https://www.dimdi.de/static/downloads/deutsch/hamburger-manual-nov2004.pdf>
Bouwstra H, Smit EB, Wattel EM et al. (2018) Measurement Properties of the Barthel Index in Geriatric Rehabilitation. JAMDA 20:420-425 DOI:<https://doi.org/10.1016/j.jamda.2018.09.033>
Hartigan I, O'Mahony D (2011) The Barthel Index: comparing inter-rater reliability between nurses and doctors in an older adult rehabilitation unit. Appl Nurs Res 24:e1-7
Hopman-Rock M, van Hirtum H, de Vreede P, Freiburger E (2019) Activities of daily living in older community-dwelling persons: a systematic review of psychometric properties of instruments. Aging Clin Exp Res. 31:917-925 DOI: 10.1007/s40520-018-1034-6
Korner-Bitensky N, Wood-Dauphinee S (1995) Barthel Index Information elicited over the telephone. Is it reliable? Am J Phys Med Rehabil 74:9-18 DOI: 10.1097/00002060-199501000-00003
Krupp S, Kasper J, Gollmer C et al. (2025) Operationalisierung der Bestimmung des Barthel-Index mittels „Barthel plus“. Z Gerontol Geriat DOI: 10.1007/s00391-025-02436-2
Lübke N, Grassl A, Kundy M et al. (2001) Hamburger Einstufungsmanual zum Barthel-Index. Geriatrie Journal 1/2:41-46
Lübke N, Meinck M, von Renteln-Kruse W (2004) Der Barthel-Index in der Geriatrie. Eine Kontextanalyse zum Hamburger Einstufungsmanual. Z Gerontol Geriat 37:316–326
Mahoney FI, Barthel DW (1965) Functional evaluation: the Barthel Index. Maryland State Med J 14:61-65
Sainsbury A, Seebass G, Bansal A, Young JB (2005) Reliability of the Barthel Index when used with older people. Age Ageing 34:228-232 DOI: 10.1093/ageing/afi063

Tab. 2.2.2: Dimension Selbsthilfefähigkeit

Name des Instrumentes:	Instrumental Activities of Daily Living	
Kürzel	IADL	
Mittlerer Zeitaufwand:		
Für Patient*in 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 12 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll

Durchführung durch:

Pflegekraft, Ergotherapeut*in, Sozialdienst, Ärzt*in, spezifisch geschultes Fachpersonal

Geeignet für Schweregrad:

x leicht/Früherkennung	x mittel	O schwer
------------------------	----------	----------

Gütekriterien:

Chen (2023): Systematischer Review, Metaanalyse von 21 Studien mit Erhebung der IADL bei >60-Jährigen präoperativ (n = 5690). Etwa die Hälfte der Teilnehmenden wies präoperative Defizite in den IADL auf, diese hatte ein mehr als doppelt so hohes Risiko für ein postoperatives Delir (49,9% versus 24,4%, Odds Ratio 2,26).

Dufournet (2021): Studienpopulation n = 2391, 79,9 J, 61,7% F, Gedächtnisambulanz. 68,8% der Männer und 17,7% der Frauen hatten mindestens eine der acht Aktivitäten niemals ausgeführt. Durch Adaptation des Summenscores soll deutlicher werden, wo echte Defizite liegen, gleichzeitig vermindert sich der Gender-Einfluss (Mittelwert bei Frauen ohne Adaptation 4,72, mit 4,82, Männer 3,49 bzw. 4,26).

Vergara (2012): Studienpopulation für Validierungsstudie der spanischen Übersetzung, n = 1965 (934 mit Hüft-, 1031 mit Handgelenksfraktur in 7 Notaufnahmen, 80,0 ± 8,0 J, 84,5% F, Barthel-Index 89,4 ± 19,5. Cronbachs Alpha 0,94, Faktorenladungen 0,67 bis 0,90. Für die Änderungssensitivität wurden n = 113 mit Verschlechterung im Barthel-Index 6 Monate nach Fraktur im Vergleich zu vorher betrachtet: Standardized Effect Size 0,79, Standardized Response Mean 0,84 (n = 485 ohne Änderung des Barthel-Index: 0,31 bzw. 0,38).

Edwards (1990): Studienpopulation n = 30 geriatrische Patient*innen. Interrater-Reliabilität 0,99, Retest-Reliabilität 0,93.

Begründung der Aufnahme in die Leitlinie:

Von der AGAST empfohlenes, international breit eingesetztes Instrument, dessen Gütekriterien bezüglich der „instrumental activities of daily living“ bislang von keinem anderen im geriatrischen Setting nachweislich übertroffen werden.

Limitationen / Forschungsbedarf:

Durch dichotome Bewertung (0 - 1) geht die differenzierte Beschreibung der angesprochenen Bereiche verloren. Es sind Varianten dieser Punktvergabe im Umlauf, die versuchen, diese Schwäche auszugleichen. Es besteht weiterer Forschungsbedarf im deutschsprachigen Raum.

Literatur:

Chen A, An E, Yan E et al. (2023) Incidence of preoperative instrumental activities of daily living (IADL) dependence and adverse outcomes in older surgical patients: A systematic review and meta-analysis. J Clin Anesth 89:111151. DOI :10.1016/j.jclinane.2023.111151

Dufournet M, Moutet C, Achi S et al. (2021) Proposition of a corrected measure of the Lawton instrumental activities of daily living score. BMC Geriatrics 21:39
<https://doi.org/10.1186/s12877-020-01995-w>

Edwards MM (1990) The reliability and validity of self-report activities of daily living scales. Canadian Journal of Occupational Therapy 57(5):273-278
<https://doi.org/10.1177%2F000841749005700507>

Hopkins RO, Suchyta MR, Kamdar BB et al. (2017) Instrumental activities of daily living after critical illness: a systematic review. Ann Am Thorac Soc 14(8):1332–1343

Lawton M, Brody E (1969) Assessment of older people: self-maintaining and instrumental activities

of daily living. Gerontologist 9:179-186 https://doi.org/10.1093/geront/9.3_Part_1.179
Vergara I, Bilbao A, Orive M et al. (2012) Validation of the Spanish version of the Lawton IADL Scale for its application in elderly people. Health and Quality of Life Outcomes 10:130. DOI:10.1186/1477-7525-10-130

Tab. 2.2.3: Dimension Selbsthilfefähigkeit

Name des Instrumentes:	Geldzähltest nach Nikolaus (Timed Test of Money Counting)		
Kürzel:	TTMC		
Mittlerer Zeitaufwand: Für Patient*in 3 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 5 min			
Lernaufwand für Untersucher*in:			
☒ leicht		☐ mittel	☐ anspruchsvoll
Durchführung durch: Ergotherapeut*in, Ärzt*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung		☒ mittel	☐ schwer
Gütekriterien: Zeyfang (2012): Studienpopulation n = 73, 77,3 ± 7,1 J, Diabetiker*innen, die vor der Diabetes-Schulung den TTMC absolvierten und 3 Monate nach der Schulung in Bezug auf die selbstständige Insulingabe untersucht wurden. Bei einem Cut-off von <46/≥46 Sekunden Sensitivität von 69%, positiver prädiktiver Wert 0,79 bezüglich der Fähigkeit zur selbstständigen Insulininjektion. <u>Nikolaus (1995a)</u> : n = 78, 80,1 (66 – 91) J, konsekutive Patient*innen ohne schwere Demenz in stationärer Geriatrie. Interrater-Reliabilität (n = 20) 0,97-0,98 (ICC). Retest-Reliabilität 0,82 (ICC). Nikolaus (1995b): n = 183, davon 78 poststationäre (s. o.), 105 nicht-institutionalisiert wohnende ambulante Patient*innen, 78,9 (66 – 95) J. Bei Cut-off 70/>70 s positiver prädiktiver Wert 0,85 für das Entwickeln zusätzlichen Pflegebedarfs innerhalb von 18 Monaten.			
Begründung der Aufnahme in die Leitlinie: Von der AGAST empfohlener Performance-Test, der eine alltagsrelevante Aktivität überprüft.			
Limitationen / Forschungsbedarf: Sofern eine Fähigkeitsstörung festgestellt wird, sollte im zweiten Schritt die Ursache eruiert werden, da das Testergebnis von den Funktionen in mehreren Dimensionen (Kognition, Visus, Feinmotorik) abhängig ist. Patient*innen mit hohem Pflegebedarf oder fortgeschrittener Demenz waren in der Validierungsstudie ausgeschlossen. Keine Angaben zur Änderungssensitivität.			

Literatur:

Nikolaus T, Bach M, Specht-Leible N et al. (1995a) The Timed Test of Money Counting: a short physical performance test for manual dexterity and cognitive capacity. Age Ageing 24:257-258

Nikolaus T, Bach M, Oster P, Schlierf G (1995b) The Timed Test of Money Counting. A simple method of recognizing geriatric patients at risk for increased health care. Aging Clin Exp Res 7:179-183

Zeyfang A, Berndt S, Aurnhammer G et al. (2012) A short easy test can detect ability for autonomous insulin injection by the elderly with diabetes mellitus. J Am Med Dir Assoc 13:81

2.3 Mobilität / Motorik

Ein umfassendes Assessment der Mobilität beinhaltet gemäß Beauchamp et al. [Beauchamp 2023] eine Betrachtung der „wahrgenommenen Mobilität“ (perceived mobility; Selbsteinschätzung der Mobilitäts-Möglichkeiten), „realen Mobilität“ (actual mobility; Häufigkeit und Dauer der Ausführung, über Fragebögen oder digital ermittelt) und der lokomotorischen Kapazität (locomotor capacity for mobility, Messung von Funktionen im Rahmen standardisierter Performanz). Bis auf den Parker Mobility Score und die Falls Efficacy Scale-International gehören alle in diesem Kapitel präsentierten Instrumente der letztgenannten Kategorie an. Digital erhobene Daten aus der Alltags-Beobachtung werden die Rolle der realen Mobilität im Assessment künftig wahrscheinlich aufwerten. Auch Umweltfaktoren modifizieren die drei Aspekte der Mobilität.

Das von Parker und Palmer 1993 zur Prognoseabschätzung veröffentlichte Bewertungssystem der Mobilität vor einem Akutereignis ist weiterhin auch unter dem Namen „New Mobility Score“ bekannt, sollte jedoch als „**Parker Mobility Score**“ (Tabelle 2.3.1) bezeichnet werden. Der Einsatz empfiehlt sich, sofern sich zuverlässige anamnestische Angaben erhalten lassen, im Rahmen des sozialen Assessments im Unterpunkt „häusliche und außerhäusliche Aktivitäten“, zu denen auch das Gehen und Einkaufen gehören. Dann ist allerdings nicht die Übermittlung des Scores im abschließenden Behandlungsbericht ausreichend, sondern (auch) die inhaltliche Beschreibung notwendig. Sinnvoll ergänzt wird diese durch die Erfassung der Versorgung mit mobilitätsfördernden Hilfsmitteln (wann wurde welches eingesetzt?) und die Fähigkeit zum Treppensteigen. Eine entsprechend erweiterte standardisierte Abfrage ist ebenso wie eine autorisierte deutsche Übersetzung des Originals verfügbar [Krupp 2017].

Der **Timed Up and Go (TUG)** (engl.: Zeitdauer für Aufstehen und Gehen) (Tabelle 2.3.2) ist der gebräuchlichste Mobilitätstest in der Geriatrie und spielt daher nicht nur für die Erhebung des aktuellen Status, sondern auch für die Langzeitbeobachtung eine wichtige Rolle. Das Ergebnis ist von Muskelkraft, Gelenkfunktion und Gleichgewicht abhängig. Der Test beurteilt die Fähigkeit zu selbstständigem Aufstehen und Gehen auf Zimmerebene. Sofern Gehhilfsmittel eingesetzt werden, sind diese zu dokumentieren. Im Setting stationärer geriatrischer Therapie ist der Anteil an Patient*innen, die den TUG in unter 15 Sekunden absolvieren und keine klinischen oder anamnestischen Hinweise auf eine Gangstörung oder ein erhöhtes Sturzrisiko zeigen, niedrig – bei allen anderen reicht ein solches der Stufe 2a entsprechendes Assessment nicht aus, sondern sollte je nach Schweregrad und vermuteter Störung durch andere Assessment-Instrumente

ergänzt oder ersetzt werden. Auch das Ergebnis „Aufgabe nicht bewältigt“ ist eine wichtige Information, sollte zum Zweck der Verlaufsbeurteilung jedoch mindestens um die differenzierte Beschreibung einer therapeutisch fokussierten lokomotorischen Fähigkeit ergänzt werden. Der TUG kann auch mit einer motorischen oder kognitiven Zusatz-Aufgabe (dual task) durchgeführt werden oder in Zukunft „Smartphone basiert“ im Selbsttest-Verfahren.

Ist ein*e Patient*in gehfähig, aber noch nicht sehr belastbar oder nicht in der Lage, aus dem Stuhl selbstständig aufzustehen, so bieten sich die verschiedenen **Gehtests** (Tabellen 2.3.3 und 2.3.4) an. Dabei ist der Einsatz von Hilfsmitteln erlaubt, soll aber dokumentiert werden. Gemessen wird die Zeit in Sekunden, die benötigt wird, um eine bestimmte Strecke zurückzulegen oder die Strecke in Metern, die in einer definierten Zeit zurückgelegt wird. Dabei ist anzugeben, ob es sich um die habituelle („wie üblich“) oder forcierte („so schnell wie sicher möglich“) Gehgeschwindigkeit handelt. Bei längeren Messungen ist die habituelle Gehgeschwindigkeit üblicher.

Die habituelle Gehgeschwindigkeit lässt Rückschlüsse auf die Mobilität im Alltag zu, z. B. die Sturzgefährdung und die Gehfähigkeit im Außenbereich einschließlich der Fähigkeit, eine Straße in der Grünphase der Ampel zu überqueren [Carvalho 2010, Kempen 2011]. Eine Geschwindigkeit von 0,8 m/s oder niedriger wird als Hinweis sowohl auf Frailty als auch auf Sarkopenie gewertet. Wenn die Proband*in mehr als 8 Sekunden benötigt, um 4 Meter zurückzulegen, so lässt es auf eine erhebliche Einschränkung im Alltag schließen. Der 4-Meter-Gehtest kann allein oder als Untertest der Short Physical Performance Battery genutzt werden. Gehtests mit Zeitnahme sind eine einfache Methode, den Schweregrad einer Gehstörung zu quantifizieren oder auch den Rehabilitationserfolg abzubilden.

Bei kurzen Teststrecken ist die Berücksichtigung zusätzlicher Strecken (mindestens je 1 m) für das Beschleunigen bzw. Abbremsen wichtig. Bei Schlaganfallpatient*innen liegen die meisten Studien für den 10-Meter-Gehtest in habituellem Tempo vor, die Gehgeschwindigkeit zeigte sich jedoch weitgehend unabhängig von der zurückgelegten Strecke, sondern wurde durch den Auftrag habituell versus forciert zu gehen sowie die Bodenbeschaffenheit beeinflusst [Cheng 2021]. Dies unterstreicht die Wichtigkeit einer hohen Standardisierung im Geriatrie-Team.

Gehtests können indirekt sowohl die kardio-pulmonale Leistungsfähigkeit [Haas 2000] als auch die funktionelle Trainingskapazität bei COPD [Johnston 2017] abbilden. Der 6- oder 2-Minuten-Gehtest sowie der 400 m-Gehtest können entweder auf dem Laufband (mit Maßnahmen zur Risikoreduktion wie bei der Fahrrad-Ergometrie) oder auf einem Korridor (Vorgabe mindestens 30 m Länge) durchgeführt werden. Wegen der Längenvorgabe sind diese Gehtests nicht überall anwendbar, für die kürzeren hingegen findet sich in der Regel ein ausreichend langer Korridor.

Die beiden Varianten des **Stuhl-Aufstehetests** (Tabelle 2.3.5) dienen dazu, Informationen zur Funktion und Kraft der unteren Extremitäten zu bekommen und spielen eine Rolle in der Diagnosestellung der Sarkopenie. Die Muskelschnellkraft (muscle power) lässt sich basierend auf dem 5-Chair Rise Test anhand von Algorithmen berechnen und spielt in der Sturzprävention eine immer wichtigere Rolle. Bei beiden Testversionen sitzen die Patient*innen auf einem Stuhl und kreuzen die Arme vor der Brust. In der 5-Chair Rise Version (**5-CRT**) wird die Zeit gemessen, die benötigt wird, um fünfmal vom Stuhl

aufzustehen. Ist dies nach 60 Sekunden nicht geschafft, gilt der Test als nicht bestanden. Dennoch sollte dokumentiert werden, wie oft es in einer Minute geschafft wurde (0-4x). Jede Modifikation der Testdurchführung muss protokolliert werden. In der 30-Sekunden Version (**30-s-CST**) wird gezählt, wie oft ein*e Teilnehmer*in in dieser Zeit vom Stuhl aufstehen kann. Beide Tests haben bei geriatrischen Krankenhauspatient*innen einen hohen Bodeneffekt.

Für nicht gehfähige Patient*innen findet in einigen Kliniken die **Esslinger Transferskala (ETS)** [Runge 1996, 2001] Verwendung. Es sind jedoch keine dezidierten Informationen zu der vom Erstautoren genannten Interrater- und Retest-Reliabilität von 0,80 und keine weiteren Literaturangaben verfügbar. Wegen schlechter Abgrenzbarkeit der Stufen H2/H3 ist eine ausreichende Validierung mit Klärung wesentlicher Fragen zu den Gütekriterien auch künftig nicht zu erwarten.

Die **Lübecker Skala der Basis-Mobilität (LSBM)** (Tabelle 2.3.6) wurde als Alternative für Patient*innen entwickelt, die den TUG nicht bewältigen, hat sich jedoch auch bei mobileren als änderungssensitiv erwiesen. Die sieben Aufgaben bilden ausgehend von der Rückenlage den Ablauf Drehen in die Seitlage - an die Bettkante setzen - Sitzposition halten (1 min) - Umsetzen - Aufstehen - Stand (1 min) - 3 m Gehen und Rückkehr ab, wobei jedes Item gesondert bewertet wird. Die fünf Schweregrade orientieren sich an der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) [WHO 2005] (z. B. LSBM Schweregrad 1: selbstständig, aber beeinträchtigt ~ ICF Grad der Einschränkung der Funktionsfähigkeit 5-24%). Bei Patient*innen, die bereits bei der Basis-Erhebung alle Items ohne Aufsicht bewältigen, wird die zusätzliche Zeiterfassung der fünf Transfere empfohlen (**t-LSBM**).

Der **De Morton Mobility Index (DEMMI)** (Tabelle 2.3.7) wurde erstmals 2008 publiziert. Mittels Rasch Analyse wurden 15 Items identifiziert, um die Mobilität bei älteren Personen zu beurteilen. Geprüft werden einfachere bis komplexere Bewegungsabläufe in fünf Subkategorien (Bett, Stuhl, statisches Gleichgewicht, Gehen und dynamisches Gleichgewicht). Der Summenscore wird anhand einer Tabelle in den DEMMI-Score (0-100 Punkte) konvertiert. Der Test wurde im Krankenhaus/Akutspital, in der Rehabilitation und bei zu Hause lebenden Menschen untersucht und validiert, auch die deutsche Übersetzung ist validiert.

Die **Short Physical Performance Battery (SPPB)** (Tabelle 2.3.8) gewinnt im geriatrischen Assessment immer mehr an Bedeutung. Zum einen deckt die SPPB die wichtigen motorischen Bereiche Gleichgewicht, Kraft und habituelle Gehgeschwindigkeit ab, zum anderen ist diese Testbatterie in großen epidemiologischen Studien eingesetzt worden und dadurch sind genügend Datensätze zum Vergleich vorhanden. Sie benötigt keinen großen Geräteaufwand und lässt sich leicht durchführen. Die SPPB ist hoch prädiktiv für negative Gesundheitseffekte und hoch sensitiv für Veränderungen und lässt sich damit auch gut in Interventionsstudien einsetzen. Bei Personen, die keine 10 Sekunden frei mit geschlossenen Füßen („side by side“) stehen können, ist der Bodeneffekt erreicht. Bei fitten älteren Menschen gibt es einen Deckeneffekt, dann können eventuelle Verbesserungen nicht angezeigt werden. Die SPPB ist eine der wichtigsten international angewandten Testbatterien. Die European Medicine Agency (EMA) hat sie als für medizinische Studien und auch zur Erfassung von Sarkopenie und Frailty geeignet anerkannt. Seit 2021 ist auch eine elektronische Version der SPPB für

Android und Windows erhältlich.

Der **Mobilitätstest nach Tinetti** (Tabelle 2.3.9, auch **POMA I und II**) dient zur Abschätzung des Sturzrisikos bei Personen, die ohne Hilfe aufstehen können. Mit einem Punktesystem werden das Gleichgewicht und das Gangbild nach definierten Kriterien bewertet. Da die Untersuchung etwa zwanzig Minuten dauert und eine spezielle Schulung sowie Erfahrung in der Begutachtung erfordert, wird sie nur noch selten eingesetzt, meist durch Physiotherapeut*innen.

Die **Berg Balance Scale (BBS)** (Tabelle 2.3.10) wurde von K. Berg für den Pflegeheimbereich zum Testen von Gang und Gleichgewicht entwickelt. In Deutschland wird dieser Test häufig von Physiotherapeut*innen im ambulanten Bereich eingesetzt. Hierfür fehlen aber wesentliche Validitäts- und Reliabilitätsergebnisse. Auch ist die Prädiktion der BBS abhängig vom funktionellen Level der Proband*innen. Insgesamt ist die BBS am ehesten für geriatrische Patient*innen mit Schlaganfall geeignet [Downs 2015].

Während leichte Sturzbedenken über erhöhte Vorsicht einen „protektiven“ Charakter haben können, wird die Angst vor Stürzen in Leitlinien und einem systematischen Review zum Thema Sturz als ein unabhängiger Risikofaktor für Stürze genannt. Sie tritt bei geriatrischen Patient*innen auch häufig ohne vorheriges Sturzereignis auf und kann zu negativer Verhaltensänderung mit weiteren negativen Gesundheitseffekten führen wie z. B. einer Reduzierung der körperlichen Aktivität und nachfolgend einem schnelleren Verlust von motorischer Funktion. Bei Sturzpatient*innen sollte daran gedacht werden, den Grad der Angst vor Stürzen mit zu erheben. Ein Kommentar von Ellmers et al. (2023) geht auf die Schwierigkeit der Wortwahl zur Erhebung von Sturzbedenken ein. Viele Vorarbeiten hatten bereits gezeigt, dass man bei der Befragung das Wort „Bedenken“ statt „Angst vor Stürzen“ wählen sollte. Diese Formulierung wurde in die World Falls Guidelines [Montero-Odasso 2022] übernommen.

Die **Falls Efficacy Scale-International (FES-I)** (Tabelle 2.3.11) mit 16 Items wurde in einer internationalen Expertengruppe entwickelt und ist seitdem samt ihrer 7 Items enthaltenden Kurzform (Short FES-I) (Tabelle 2.3.12) in der Sturzforschung ein wichtiges Instrument, um Sturzbedenken bzw. die psychologischen Konsequenzen von Stürzen zu erfassen. Es sind Übersetzungen in zahlreiche Sprachen vorhanden [<https://sites.manchester.ac.uk/fes-i/>] sowie eine lange (30 Items, Icon FES) und kurze (10 Items, Short Icon FES) ikonografische Version, die die Anwendung besonders bei kognitiv beeinträchtigten Personen erleichtern soll.

Um den Einfluss von Sturzbedenken auf die körperliche Aktivität im Sinne von Vermeidungsverhalten zu erfassen, wurde die FES-I Avoidance Behavior (FES-IAB) [Dorresteijn 2011; Kruisbrink 2021] entwickelt. Die FES-IAB enthält die 16 FES-I Fragen, aber erfasst zusätzlich mit der Frage „Do you avoid performing this activity due to concerns about falling?“, wie oft die geäußerten Bedenken zu einem Vermeidungsverhalten führen. Das gleiche Prinzip gilt für die auf der Short FES-I beruhende Short FES-IAB.

Gemäß der Einteilung in der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit [WHO 2005] gehört zur Mobilität nicht nur, die Körperposition zu ändern und aufrecht zu erhalten, zu gehen und sich fortzubewegen,

sondern z. B. auch, Gegenstände zu tragen, zu bewegen und zu handhaben. Daher werden in diesem Abschnitt auch Instrumente beschrieben, die sich auf Aspekte des Handeinsatzes beziehen. Diesbezügliche Beeinträchtigungen werden oft nicht spontan geäußert, so dass die Frage danach Bestandteil der Aufnahmeanamnese sein sollte (z. B. „Haben Sie Schwierigkeiten, weil eine Ihrer Hände (oder beide) kraftloser oder ungeschickter geworden ist? Haben Sie Schwierigkeiten damit, die Hand dort hinzubekommen, wo Sie sie brauchen?“). Mindestens im Fall von Beeinträchtigungen sollten sowohl die Handkraft und Grobmotorik als auch die Feinmotorik überprüft werden.

Alterstypische Erkrankungen (z. B. Humerusfraktur, Arthrose oder Impingement-Syndrom der Schulter, Schlaganfall) verursachen häufig eine die Selbstversorgung beeinträchtigende Einschränkung des Greifraumes.

Der **8-Punkte-Greifraum-Test (8P-GRT)** (Tabelle 2.3.13) überprüft seitengetrennt die Fähigkeit, die für die Basis-Aktivitäten des täglichen Lebens wichtigsten Handpositionen einzunehmen. Sofern therapierelevante Defizite festgestellt werden, ist für die Stufe 2b des Assessments deren Ausprägung näher zu beschreiben (z. B. mit Anwendung der Neutral-Null-Methode im Bereich der betroffenen Gelenke).

Der **Handkraft** (Tabelle 2.3.14) wird eine positive Korrelation mit der Gesamtkörperkraft und Gehfähigkeit [Beseler 2014] und eine negative Korrelation mit dem Sturz- und Frakturrisiko sowie der Gesamtsterblichkeit zugeschrieben [Huemer 2023, Reuter 2011, Bohannon 2015, 2019]. Die Handkraftmessung kann überall eingesetzt werden, im ambulanten, stationären klinischen Bereich oder in der Pflegeeinrichtung. Die Durchführung sollte standardisiert erfolgen: Sitzende Position, die Proband*in legt den Unterarm auf dem Oberschenkel ab oder den Oberarm an die Körperseite an und hat den Ellenbogen im 90°-Winkel gebeugt, der Unterarm befindet sich in Neutralstellung, das Handgelenk in 0-30 Grad Extension. Das Dynamometer bzw. der Ballon des Vigorimeters wird in der Hand gehalten. Daraufhin werden mit jeder Hand dreimal mit Intervallen von mindestens 30 Sekunden möglichst kräftige Faustschlüsse durchgeführt, wobei je nach Autor*in der Durchschnitt oder der höchste Wert in die Auswertung einbezogen wird [Roberts 2011]. Logischer erscheint die Verwendung des höchsten Wertes, da „falsch hohe“ Werte kaum produziert werden können, während eine ungünstige Handhabung mit falsch niedrigem Wert vorkommt und ggf. in den Mittelwert einfließen würde. Zusätzlich kann die Handkraftmessung auch für den Frailty Phänotyp nach Fried herangezogen werden, um das Kriterium der Schwäche auszuschließen bzw. nachzuweisen. Die European Working Group on Sarcopenia in Older People (EWGSOP) legte unter Pooling britischer Daten dafür ein Unterschreiten von 27 (Männer) bzw. 16 kg (Frauen) fest [Cruz-Jentoft 2019], jedoch sind die Normwerte der Handkraft nicht nur vom Geschlecht, sondern u. a. auch vom Alter und der Ethnie abhängig. Für Vigorimeter gelten andere Normwerte.

Feinmotorische Aktivitäten setzen die Funktionsfähigkeit vieler unterschiedlicher struktureller Systeme voraus, z. B. Beweglichkeit des Gelenksystems, muskuläre Kraft und Kraftdosierung, Auge-Hand-Koordination, Tastsinn usw. Bereits eine Störung in nur einem dieser Systeme führt häufig im Alltag alter Menschen zu Beeinträchtigungen.

Zur quantifizierenden Abbildung des Funktionsdefizits anhand des im Alltag wichtigsten Feingriffs steht u. a. der **20-Cents-Test (20-C-T)** (Tabelle 2.3.15) zur Verfügung. Im Unterschied zum Geldzähltest nach Nikolaus wurde der Test so konzipiert, dass auch

Patient*innen mit mittelgradiger kognitiver oder visueller Beeinträchtigung ein normales Ergebnis erzielen können.

Tab. 2.3.1: Dimension Mobilität

Name des Instrumentes:	Parker Mobility Score (New Mobility Score)		
Kürzel:	PMS		
Mittlerer Zeitaufwand:			
Für Patient*in 1 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 2 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Ärzt*in, Physiotherapeut*in, Ergotherapeut*in, Pflegekraft, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien:			
Kusen (2021) Studienpopulation n = 649 ab 70 J mit operativer Versorgung einer Hüftfraktur in 2 Traumazentren. Nach dem Alter hatte der PMS den zweitstärksten Effekt auf den Entlassungsweg, positiver prädiktiver Wert hinsichtlich einer Verlegung in eine Rehabilitationsklinik 0,91.			
Kristensen (2008): Studienpopulation n = 48 konsekutive Patient*innen mit akuter Femurfraktur, davon 40 nicht institutionalisiert, Median 84 (range interquartil 76 – 89) J. Interrater-Reliabilität (ICC) 0,98, Konfidenzintervall 0,96-0,99, Standardfehler des Mittelwerts 0,42, bei Erheben durch 2 Physiotherapeut*innen. Patient*innen, die unterschiedlich bewertet wurden, waren signifikant häufiger kognitiv beeinträchtigt.			
Parker (1993): Studienpopulation n = 882 Patient*innen mit proximaler Femurfraktur, befragt bei Krankenhausaufnahme. Ca. die Hälfte lag bei einem Score von <5 (0 = nicht gehfähig, 9 = Maximalwert), negativer prädiktiver Wert für diese bezüglich Versterben innerhalb eines Jahres 0,81, positiver 0,51, Sensitivität 73%, Spezifität 62%.			
Begründung der Aufnahme in die Leitlinie:			
Anwendung auch zur standardisierten Anamnese der Mobilität vor einem Akutereignis, Ergebnis prognostisch relevant. Zur (ggf. telefonischen) poststationären Kontrolle geeignet.			
Limitationen / Forschungsbedarf:			
Einsatz nur sinnvoll, wenn die anamnestischen Angaben zuverlässig sind.			
Literatur:			
Garg A, Kamboj P, Sharma PK et al. (2022) Evaluation of functional outcome and comparison of three different surgical modalities for management of intertrochanteric fractures in elderly population. Int J Burns Trauma 12(1):13–22.			
Gatot C, Tan ESE, Liow MHL et al. (2021) Higher Charlson Comorbidity Index increases 90-day			

readmission rate with poorer functional outcomes in surgically treated hip fracture patients. Geriatric Orthopaedic Surgery & Rehabilitation 12: 1–10. DOI: 10.1177/21514593211036252

Karademir G, Bilgin Y, Demirel M et al. (2022) Effect of treatment modality on mobility and quality of life in unstable intertrochanteric fractures. Acibadem Univ. Sağlık Bilim. Derg. 13 (4) <https://doi.org/10.31067/acusaglik.1143274>

Kristensen MT, Bandholm T, Foss NB et al. (2008) High inter-tester reliability of the New Mobility Score in patients with hip fracture. J Rehabil Med 40:589–591

Kusen JQ, van der Naald N, van Overeem L et al. (2021) Is the Parker Mobility Score in the older patient with a traumatic hip fracture associated with discharge disposition after surgery? A retrospective cohort study. Eur J Trauma Emerg Surg DOI: 10.1007/s00068-021-01712-1

Deutsche Version in:

Krupp S (2017) Geriatrisches Assessment. In: Willkomm M (Hrsg) Praktische Geriatrie. 2. Auflage. Georg Thieme Verlag, Stuttgart, S.41

Parker MJ, Palmer CR (1993) A new mobility score for predicting mortality after hip fracture. J Bone Joint Surg Br 75:797–798

Voeten SC, Nijmeijer WS, Vermeer M et al. (2020) Validation of the Fracture Mobility Score against the Parker Mobility Score in Hip Fracture Patients. Injury 51:395-399
DOI: 10.1016/j.injury.2019.10.035

Wiedl A, Förch S, Fenwick A et al. (2022) Orthogeriatric co-management: differences in outcome between major and minor fractures. Eur J Trauma Emerg Surg 48(4):2953–2966. DOI: 10.1007/s00068-022-01974-3

Tab. 2.3.2: Dimension Mobilität

Name des Instrumentes:	Timed Up and Go		
Kürzel:	TUG		
Mittlerer Zeitaufwand:			
Für Patient*in 3 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 5 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Physiotherapeut*in, Ergotherapeut*in, Ärzt*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien:			
<u>Podsiadlo, Richardson (1991):</u> Studienpopulation n = 60, Tagesklinik, 79,5 J. Kriterium-basierte Validität: hohe Korrelation zum logarithmisch transformierten Score der Berg Balance Scale (r = -0,81), Barthel Index (r = -0,78) und Gehgeschwindigkeit (r = -0,61).			
Prognostische Validität: Stürze: Bei zu Hause lebenden älteren Menschen Cut-off mit der besten Genauigkeit für die Risikoabschätzung der Sturzgefahr (Vorhersageraten zwischen 80 und 90%) bei weniger als 12 Sekunden [Lusardi 2017]. Die World Guidelines for Falls			

Prevention [Montero-Odasso 2022] schlagen den Cut-off 15/>15 s vor.
Retest-Reliabilität: Exzellente ICC zwischen 0,96 und 0,99 je nach Population [Flansbier 2005].

Interrater-Reliabilität: ICC im Mittel 0,91 [Rydwik 2011].

Boden/Decken-Effekte: große Bodeneffekte, da für immobile Patient*innen nicht durchführbar.

Änderungssensitivität: sehr variabel, weitere Untersuchungen notwendig.

Minimaler klinisch detektierbarer Unterschied 2-4 Sekunden je nach Population [Rydwik 2011]

In einem Review [Beck Jepsen 2022] wird der TUG wegen inkonsistenter Studienergebnisse nicht mehr zur Sturzprädiktion empfohlen, jedoch in den World Guidelines for Falls Prevention [Montero-Odasso 2022] als Alternative zur Erhebung der Gehgeschwindigkeit genannt. Die Centers for Disease Control and Intervention (CDC) nennen den TUG als eine von drei Methoden im Algorithmus gemäß STEADY (Stopping Elderly Accidents, Deaths & Injuries), durchzuführen bei anamnestischem Hinweis auf erhöhtes Sturzrisiko ab 65-jähriger nicht-institutionalisiert lebender Personen.

Begründung der Aufnahme in die Leitlinie:

Bereits von der AGAST empfohlenes, international breit eingesetztes Instrument des CGA mit günstigem Aufwand/Nutzen-Verhältnis bei erhaltener Funktion des Aufstehens von einem Stuhl.

Limitationen / Forschungsbedarf:

Forschungsbedarf besteht hinsichtlich der Validität App-basierter Selbsttests.

Literatur:

Barry E, Galvin R, Keogh C et al. (2014) Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. BMC geriatrics 14:14

Beck Jepsen D, Robinson K, Ogliaeri G et al. (2022) Predicting Falls in Older Adults: An Umbrella Review of Instruments Assessing Gait, Balance, and Functional Mobility. BMC Geriatrics 22:615 <https://doi.org/10.1186/s12877-022-03271-5>

Bohannon RW (2006) Reference values for the timed up and go test: a descriptive meta-analysis. J Geriatr Phys Ther 29(2): 64-68. DOI: 10.1519/00139143-200608000-00004

Böttinger MJ, Mellone S, Klenk J et al. (2025) A Smartphone-Based Timed Up and Go Test Self-Assessment for Older Adults: Validity and Reliability Study. JMIR Aging 8. DOI:10.2196/67322

Flansbier UB, Holmbäck AM, Downham D et al. (2005) Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med 37:75-82

Long JW, Cai TP, Huang XY et al. (2020) Reference value for the TUGT in healthy older people: A systematic review and meta-analysis. Geriatric Nursing 41(3): 325-330.

Lusardi MM, Fritz S, Middleton A et al. (2017) Determining risk of falls in community dwelling older adults: A systematic review and meta-analysis using posttest probability. J Geriatr Phys Ther. 40:1-36

Mayhew AJ, So HY, Ma J et al. (2023) Normative values for grip strength, gait speed, timed up and go, single leg balance, and chair rise derived from the Canadian longitudinal study on ageing. Age Ageing 52(4). DOI:10.1093/ageing/afad054

Montero-Odasso M, van der Velde N, Martin FC (2022) World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 51: 1–36 <https://doi.org/10.1093/ageing/afac205>

Podsiadlo D, Richardson S (1991) The timed-up-and-go test. A test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 39:142

Rydwik E, Bergland A, Forséén L, Frändin K (2011) Psychometric properties of Timed Up and Go

in elderly people: a systematic review. Phys Occup Ther Geriatr 14:102–125

Steffen TM, Hacker TA, Mollinger L (2002) Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. Phys Ther 82:128-137

Svinøy OE, Hilde G, Bergland A, Strand BH (2021) Timed Up and Go: Reference Values for Community-Dwelling Older Adults with and without Arthritis and Non-Communicable Diseases: The Tromsø Study. Clin Interv Aging 16:335-343. DOI: 10.2147/CIA.S294512

Tab. 2.3.3: Dimension Mobilität

Name des Instrumentes:	4-Meter-Gehtest / 10-Meter-Gehtest		
Kürzel:	4mWT/ 10mWT		
Mittlerer Zeitaufwand: Für Patient*in 1 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 3 min			
Lernaufwand für Untersucher*in:			
☒ leicht		☐ mittel	
		☐ anspruchsvoll	
Durchführung durch: Physiotherapeut*in, Ärzt*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung		☒ mittel	
		☒ schwer	
Gütekriterien: Cabanas-Valdez (2023): ambulante Schlaganfall-Patient*innen (58,6 ± 11,3 J) absolvierten u. a. den 4m-Gehtest in nicht forcierter Geschwindigkeit. Retest-Reliabilität (ICC) 0,99, Korrelation mit 10m-Gehtest r = 0,96. Tilson (2010): Bei Schlaganfallpatient*innen ist eine Verbesserung der Gehgeschwindigkeit um 0,16 m/s klinisch relevant. Fulk (2008): Es besteht ein Zusammenhang zwischen dem 4mWT und der Mortalität und der Wahrscheinlichkeit, in ein Pflegeheim eingewiesen zu werden (siehe Short Physical Performance Battery). Es besteht ein Zusammenhang (r=0,89) zwischen dem 10mWT und dem 6-min-Gehtest.			
Begründung der Aufnahme in die Leitlinie: Die Gehgeschwindigkeit ist sowohl alltagsrelevant als auch ein vielfältiger Prognosefaktor und spielt eine Rolle in der Sarkopenie-Diagnostik.			
Literatur: Bohannon RW, Wang Y-C (2019) Four-meter Gait Speed: Normative Values and Reliability Determined for Adults Participating in the NIH Toolbox Study. Arch Phys Med Rehabil 100(3):509–513 Cabanas-Valdes R, Garcia-Rueda L, Salgueiro C et al. (2023) Assessment of the 4-meter walk test test-retest reliability and concurrent validity and its correlation with the five sit-to-stand test in chronic ambulatory stroke survivors. Gait & Posture 101:8-13. https://doi.org/10.1016/j.gaitpost.2023.01.014			

Carvalho C, Sunnerhagen KS, Willén C (2010) Walking speed and distance in different environments of subjects in the later stage post-stroke. *Physiother Theory Pract* 26:519-527

Cheng DK, Dagenais M, Alsbury-Nealy K et al. (2021). Distance-limited walk tests post-stroke: A systematic review of measurement properties. *NeuroRehabilitation* 48(4):413–439

Cruz-Jentoft AJ, Bahat G, Bauer J et al. (2019) Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing*, 48:16-31. DOI:10.1093/ageing/afy169

Fulk GD, Echternach JL, Nof L, O'Sullivan S (2008) Clinometric properties of the six-minute walk test in individuals undergoing rehabilitation poststroke. *Physiother Theory Pract* 24:195-204

Jepsen B, Robinson K, Ogliari G et al. (2022) Predicting Falls in Older Adults: An Umbrella Review of Instruments Assessing Gait, Balance, and Functional Mobility. *BMC Geriatrics* 22:615 <https://doi.org/10.1186/s12877-022-03271-5>

Kempen JC, de Groot V, Knol DL et al. (2011) Community walking can be assessed using a 10-metre timed walk test. *Mult Scler* 17:980-990

Krumpoch S, Lindemann U, Rappl A et al. (2021) The effect of different test protocols and walking distances on gait speed in older persons. *Aging Clin Exp Res* 33(1): 141-146.

Middleton A, Fritz SL, Lusardi M (2015) Walking speed: the functional vital sign. *J Aging Phys Act* 23(2): 314-322. DOI: 10.1123/japa.2013-0236

Tilson JK, Sullivan KJ, Cen SY et al. (2010) Locomotor Experience Applied Post Stroke (LEAPS) Investigative Team. Meaningful gait speed improvement during the first 60 days poststroke: minimal clinically important difference. *Phys Ther* 90:196-208

Tab. 2.3.4: Dimension Mobilität

Name des Instrumentes:	2-Minuten-Gehtest / 6-Minuten-Gehtest	
Kürzel	2-min-WT/ 6-min-WT	
Mittlerer Zeitaufwand:		
Für Patient*in 3 min/7 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 5 min/9 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Physiotherapeut*in, Ärzt*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien:		
Willi (2022): n = 50, 53 ± 16 J, 34% F, Rückenmarksverletzung. Bei forciertem Gang („as far as possible, but safely“) korrelierte der 2-min-WT mit dem 6-min-WT 0,97 (bei Patient*innen mit Gehgeschwindigkeit unterhalb des Medians von 0.9 m/s mit der im 10m-WT 0,93, bei den schnelleren Patient*innen 0,89). Retest-Reliabilität 2-min-WT (ICC) 0,98. Die Autor*innen favorisieren den 2-min-WT insbesondere bei leistungsschwachen Patient*innen, allerdings waren Patient*innen, die im 6-min-WT unter 60 m zurückgelegt hatten, von der Teilnahme ausgeschlossen.		

Regan (2020): n = 66, 66 ± 12 J, 33% F, im Mittel 12 J nach Insult, ambulant. 6-min-WT 274,1 ± 152,5 m. Bei niedriger Gehgeschwindigkeit war das Sturzrisiko gemäß Activities Specific Balance Confidence Scale erhöht (Pretest-Wahrscheinlichkeit 59,1%, bei <331,65 m 74,3%) Interrater-Reliabilität: zwischen 0,81 [Lim 2005] und 0,95 [Green 2002]. Validität: (r=0,69) [Fulk 2008]. Responsivität: Im 6-min-WT gelten Distanzen ab 51 m als minimal detectable change [Flansbjer 2005]. Im 2-min-WT liegt die kleinste klinisch relevante Änderung bei COPD Patient*innen bei 5,5 m. Begründung für die Aufnahme in die Leitlinie: Zwei- und mehrminütige Gehtests geben Aufschluss über die Ausdauer einschließlich kardiopulmonaler Leistung. Limitationen / Forschungsbedarf: Nur sinnvoll durchführbar, wenn Gehfähigkeit gegeben ist. Literatur: ATS statement: guidelines for the six-minute walk test (2002) Am J Respir Crit Care Med 166:111-117 Flansbjer UB, Holmbäck AM, Downham D et al. (2005) Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med 37:75-82 Fulk GD, Echternach JL, Nof L, O'Sullivan S (2008) Clinometric properties of the six-minute walk test in individuals undergoing rehabilitation poststroke. Physiother Theory Pract 24:195-204 Graham JE, Ostir GV, Kuo YF et al. (2008) Relationship between test methodology and mean velocity in timed walk tests: a review. Arch Phys Med Rehabil 89:865-872 Green J, Forster A, Young J (2002) Reliability of gait speed measured by a timed walking test in patients one year after stroke. Clin Rehabil 16:306-314 Johnston KN, Potter AJ, Phillips AC (2017). Minimal important difference and responsiveness of 2-minute walk test performance in people with COPD undergoing pulmonary rehabilitation. International journal of chronic obstructive pulmonary disease. 12:2849–2857 Lenssen AF, Wijnen LC, Vankan DG et al. (2010) Six-minute walking test done in a hallway or on a treadmill: how close do the two methods agree? Eur J Cardiovasc Prev Rehabil 17:713-717 Lim LI, van Wegen EE, de Goede CJ et al. (2005) Measuring gait and gait-related activities in Parkinson's patients own home environment: a reliability, responsiveness and feasibility study. Parkinsonism Relat Disord 11:19-24 Regan E, Middleton A, Stewart JC et al. (2020) The six-minute walk test as a fall risk screening tool in community programs for persons with stroke: a cross-sectional analysis. Topics in stroke rehabilitation 27(2):118–126 Willi R, Widmer M, Merz N et al. (2023) Validity and reliability of the 2-minute walk test in individuals with spinal cord injury. Spinal Cord 61:15–21. https://doi.org/10.1038/s41393-022-00847-1
--

Tab. 2.3.5: Dimension Mobilität

Name des Instrumentes:	5 Chair Rise Test/30 s Chair Stand Test
Kürzel:	5-CRT/30-s CST
Mittlerer Zeitaufwand:	
Für Patient*in 2 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 4min	

Lernaufwand für Untersucher*in:

☒ leicht	☐ mittel	☐ anspruchsvoll
---	------------------------------	-------------------------------------

Durchführung durch:

Physiotherapeut*in, Ärzt*in, Ergotherapeut*in, spezifisch geschultes Fachpersonal

Geeignet für Schweregrad:

☒ leicht/Früherkennung	☒ mittel	☐ schwer
---	---	------------------------------

Gütekriterien:

Gute prädiktive Validität zur Beinkraft [Jones 1999].
Gute Reliabilität des 5-CRT (ICC 0,89-1,0) [Ng 2015] und 30-s CST (ICC 0,79-0,93) [Jones 1999], auch bei Schlaganfallpatient*innen (ICC 0,90-0,99).
Gute Änderungssensitivität des 30-s-CST mit Effektstärke von 0,83 zwischen aktiven und nicht aktiven älteren Menschen [Hansen 2017].
Die European Working Group on Sarcopenia in Older People (EWGSOP) empfiehlt den 5-CRT mit einem Cut-off von 15/>15 s neben der Erhebung der Handkraft zur Feststellung von Muskelschwäche [Cruz-Jentoft 2019].
Beim 30-s-CST wird ein Cut-off von >8/≤8 berichtet, um eine gute von einer geringen körperlichen Leistungsfähigkeit zu unterscheiden [Hansen 2017].
Es liegen Altersgruppen gemäße Cut-off Werte für diesen Test vor [Mayhew 2023].
Untersucht worden ist auch der Effekt der Beinstellung und Stuhlhöhe auf das Ergebnis beim Stuhlaufstehen [Reider 2016].

Begründung für die Aufnahme in die Leitlinie:

Die Funktion des Aufzustehens ist sowohl alltagsrelevant als auch repräsentativ für die Muskelkraft. Sie spielt eine Rolle in der Sarkopenie-Diagnostik.

Limitationen / Forschungsbedarf:

Nur sinnvoll durchführbar, wenn wenigstens selbstständiger Transfer möglich ist.

Literatur:

Albalwi AA, Alharbi AA (2023) Optimal procedure and characteristics in using five times sit to stand test among older adults: A systematic review. Medicine (Baltimore) 102(26):e34160. DOI: 10.1097/MD.00000000000034160
Bohannon RW (2012) Measurement of Sit-to-Stand among older adults. Topics in Geriatric Rehabilitation 28(1):11-16. DOI:10.1097/TGR.0b013e31823415fa
Coelho-Junior HJ, Marzetti E, Picca A et al. (2024) Sex- and age-specific normative values of lower extremity muscle power in Italian community-dwellers. J Cachexia Sarcopenia Muscle 15(1): 45-54
Cruz-Jentoft AJ, Bahat G, Bauer J et al. (2019) Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing, 48:16-31. DOI:10.1093/ageing/afy169
Hansen Bruun I, Mogensen CB, Norgaard B et al. (2019) Validity and responsiveness to change of the 30-Second Chair-Stand Test in older adults admitted to an emergency department. J Geriatr Phys Ther 42(4):265-274. DOI:10.1519/JPT.0000000000000166
Jones CJ, Rikli RE, Beam WC (1999) A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 70:113-119
Mayhew AJ, So HY, Ma J et al. (2023) Normative values for grip strength, gait speed, timed up and go, single leg balance, and chair rise derived from the Canadian longitudinal study on ageing. Age Ageing 52(4). DOI:10.1093/ageing/afad054
Ng SS, Kwong PW, Chau MS et al. (2015) Effect of arm position and foot placement on the five times sit-to-stand test completion times of female adults older

than 50 years of age. J Phys Ther Sci 27:1755-1759

Reider N, Gaul C (2016) Fall risk screening in the elderly: A comparison of the minimal chair height standing ability test and 5-repetition sit-to-stand test. Arch Gerontol Geriatr 65:133-139

Rikli RE, Jones CJ (1999) Functional Fitness Normative Scores for Community-Residing Older Adults, Ages 60-94. J Aging Phys Act 7:162-181

Tab. 2.3.6: Dimension Mobilität

Name des Instrumentes:	Lübecker Skala der Basis-Mobilität	
Kürzel:	LSBM (mit Zeiterfassung: t-LSBM)	
Mittlerer Zeitaufwand:		
Für Patient*in 7 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 10 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Physiotherapeut*in, Pflegekraft, Ärzt*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☐ leicht/Früherkennung	☒ mittel	☒ schwer
Gütekriterien:		
Krupp (2025): 85 akutgeriatrische stationäre Patient*innen, die den TUG bei Aufnahme bewältigt hatten, 82 ± 6 J, 66% F, Barthel-Index 58 ±10. Bei Intervall der Messungen von im Mittel 13 Tagen Änderungssensitivität über Cohens d für TUG 0,56, für LSBM und t-LSBM 0,77. Bei 18,8% der „TUG-fähigen“ Patient*innen Deckeneffekt der LSBM erreicht, der mittels ergänzender Zeitnahme (t-LSBM) aufgehoben wird. Krupp (2023): 77 akutgeriatrische stationäre Patient*innen, die den TUG bei Aufnahme nicht bewältigt hatten, 82 ± 8 J, 60% F, Barthel-Index 45 ± 15. Hohe Werte für Interrater-Reliabilität 0,98, Korrelation Test/Retest 0,84, interne Konsistenz über Cronbachs α 0,88, Änderungssensitivität über Effektstärke Cohens d 0,71, Korrelation mit DEMMI -0,88, kein Bodeneffekt bei Aufnahme (bei DEMMI bei 6,5%), prädiktive Validität für Bewältigung des TUG bei Entlassung -0,58 (DEMMI 0,54). Der kognitive Status (bei 36% deutliche, bei 26% leichte kognitive Defizite) korrelierte nicht signifikant mit dem Ergebnis der LSBM.		
Begründung für die Aufnahme in die Leitlinie:		
LSBM: änderungssensitiv auch bei Patient*innen, die den TUG nicht bewältigen. Bei höhergradiger kognitiver Störung einsetzbar, geringerer Bodeneffekt als der DEMMI. Zusätzliche Zeitnahme (→ t-LSBM) vermeidet Deckeneffekt bei nur leicht reduzierter Mobilität.		
Limitationen / Forschungsbedarf:		
Das Intervall zwischen Test und Retest lag bei der LSBM im Mittel bei 9 Tagen, die Retest-		

Reliabilität dürfte bei kürzerem Intervall oberhalb von 0,84 liegen.

Literatur:

Deutsche Version in:

Krupp S, Wentzel R, Balck F et al. (2023): Die Lübecker Skala der Basis-Mobilität. Assessment schwer in ihrer Mobilität beeinträchtigter geriatrischer Patienten. Z Gerontol Geriat 56(6):470-476. DOI: 10.1007/s00391-023-02220-0

Krupp S, Müller J, Kasten M, Kasper J (2025) Luebeck Scale of Basic Mobility among geriatric hospital patients: LSBM and timed LSBM. Z Gerontol Geriat DOI: 10.1007/s00391-025-02484-8

Tab. 2.3.7: Dimension Mobilität

Name des Instrumentes:	de Morton Mobility Index		
Kürzel:	DEMMI		
Mittlerer Zeitaufwand:			
Für Patient*in ca. 9 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 11 min			
Lernaufwand für Untersucher*in:			
O leicht	x mittel	O anspruchsvoll	
Durchführung durch:			
Physiotherapeut*in, Ärzt*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
x leicht/Früherkennung	x mittel	x schwer	
Gütekriterien:			
Søndergaard (2022): Studienpopulation n = 250, mittleres Alter 85 J, 63% F, ambulante Rehabilitation. Ein Score unter 67 sagte fehlende Fähigkeit zur Benutzung öffentlicher Verkehrsmittel voraus, ein Score unter 48 fehlende selbstständige Gehfähigkeit in der Umgebung der Wohnung. Die Änderung des DEMMI-Scores korrelierte mit der des Timed Up & Go sowie des 4-m-Gehtests -0.45. AUC für das Erkennen von Verbesserung gemäß Selbstbeurteilung/Urteil von Therapeuten 0,66/0,63. Hulsbæk (2019): 222 stationäre Patient*innen mit Hüftfraktur, im Mittel 83 J (64-100 J), 71% F, 26% aus Pflegeheimen, alle vor Fraktur gehfähig gewesen. Untersuchung 1 Tag postoperativ und zum Verlegungszeitpunkt in die stationäre Geriatrie 9 ± 5 Tage postoperativ. Bodeneffekt 31% (17% bei Patient*innen, die von Zuhause kamen, 72% bei denen aus Pflegeheimen) bei zu diesem Zeitpunkt Barthel-Index von 48 ± 32. Trøstrup (2019): Validation der dänischen Übersetzung an 235 Krankenhauspatient*innen 85 ± 7 J, 69% F, 2% aus Pflegeheimen, Barthel-Index 59 ± 24. Der DEMMI-Score korrelierte bei Aufnahme/Entlassung r = 0,68/0,70 mit dem Barthel-Index. Patient*innen, die nach Hause entlassen wurden, hatten bei Aufnahme einen mittleren DEMMI-Score von 55,6 ± 15,3 gehabt (n = 101), die mit Verlegung in ein Pflegeheim oder zur stationären			

Rehabilitation von $45,2 \pm 14,5$ ($n = 54$); Minimal Clinically Important Difference gemäß Vergleich mit fünfstufigem Gesamturteil zur Änderung der Mobilität durch Therapeuten 18,8, durch Patient*innen 10,7.

Braun (2015a): 133 stationäre Patient*innen in der Akut- ($n = 62$) und Reha-Geriatrie ($n = 71$), 80 ± 8 J (51-96 J). Bodeneffekt bei Aufnahme 9/62 (14,5%) bzw. 7/71 (9,9%), bei Entlassung aus der Akutgeriatrie 3/26 (11,5%) bzw. aus der Reha-Geriatrie 1/37 (2,7%). Bodeneffekt des Timed Up & Go akutstationär bei Aufnahme 40/62 (64,5%), bei Entlassung 11/26 (42,3%).

De Morton (2008): 106 Krankenhauspatienten, davon 1,9% aus Pflegeheimen, 44,6% vor Aufnahme ohne Hilfsmittel gehfähig gewesen, 81 ± 7 J, 47% F, Barthel-Index $82,5 \pm 18,8$, Bodeneffekt $<1\%$. Hohe Interrater-Reliabilität (Pearsons $r = 0,94$, im Mittel 9,5 Punkte Unterschied im Demmi-Score; Minimal Clinically Important Difference 10 Punkte im Demmi-Score, korrelierte $r = 0,91$ mit Hierarchical Assessment of Balance and Mobility (HABAM), $r = 0,68$ mit Barthel-Index; Änderungssensitivität gemäß Effektgröße 0,39.

Begründung der Aufnahme in die Leitlinie:

Das gut standardisierte Instrument erlaubt eine Beurteilung über ein großes Leistungsspektrum hinweg bei geringem Boden- und Deckeneffekt.

Limitationen / Forschungsbedarf:

Zur Charakterisierung von Mobilitätsproblemen ist die Leistung nicht nur über den Score, sondern auch Item-bezogen zu betrachten.

Literatur:

Braun T, Schulz RJ, Hoffmann M et al. (2015a) Deutsche Version des De Morton Mobility Index. Erste klinische Ergebnisse aus dem Prozess der interkulturellen Adaptation. Z Gerontol Geriatr 48 :154-163

Braun T, Schulz RJ, Reinke J et al. (2015b) Reliability and validity of the German translation of the de Morton Mobility Index (DEMMI) performed by physiotherapists in patients admitted to a sub-acute inpatient geriatric rehabilitation clinic. BMC Geriatr 15:58

Dasenbrock L, Berg T, Lurz S et al. (2016) Der De Morton Mobility Index zur Evaluation der geriatrischen Frührehabilitation. Z Gerontol Geriatr 49:398-404

De Morton NA, Jones CT, Keating JL (2008) The de Morton Mobility Index (DEMMI): An essential health index for an ageing world. Health Qual Life Outcomes 6:63

De Morton NA, Davidson M, Keating JL (2010) Validity, responsiveness and the minimal clinical important difference for the de Morton Mobility Index (DEMMI) in an older acute medical population. BMC Geriatr 10:72

De Morton NA, Davidson M, Keating JL (2011) Reliability of the de Morton Mobility Index (DEMMI) in an Older Acute Medical Population. Physiother Res Int 16:159-169

De Morton NA, Nolan J, O'Brien M. et al. (2015) Validity, responsiveness and the minimal clinical important difference for the de Morton Mobility Index (DEMMI) in an older acute medical population. Disabil Rehabil 37:1881-1887

Hulsbæk S, Larsen RF, Rosthøj S, Kristensen MT (2019) The Barthel Index and the Cumulated Ambulation Score are superior to the de Morton Mobility Index for the early assessment of outcome in patients with a hip fracture admitted to an acute geriatric ward. Disabil Rehabil 41:1351-1359. DOI: 10.1080/09638288.2018.14249516.

Søndergaard K, Petersen LE, Pedersen M et al. (2022) The responsiveness and predictive validity of the de Morton Mobility Index in geriatric rehabilitation. Disability and Rehabilitation 44(3): 478-486 DOI: 10.1080/09638288.2020.1771438

Trøstrup J, Andersen H, Kam CA et al. (2019) Assessment of Mobility in Older People Hospitalized

for Medical Illness Using de Morton Mobility Index and Cumulated Ambulation Score - Validity and Minimal Clinical Important Difference. J Geriatr Phys Ther 42(3): 153-160
DOI: 10.1519/JPT.0000000000000170

Tab. 2.3.8: Dimension Mobilität

Name des Instrumentes:	Short Physical Performance Battery		
Kürzel:	SPPB		
Mittlerer Zeitaufwand:			
Für Patient*in 12 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 16 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Physiotherapeut*in, Ärzt*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien:			
Gute prädiktive Validität für Mobilität, Beeinträchtigung der Alltagsaktivitäten, Mortalität [Guralnik 1994, 2000], Länge des Krankenhausaufenthalts [Volpato 2011] und Entlassung in ein Pflegeheim. Nicht geeignet zur Sturzprädiktion [Beck Jepsen 2022].			
Gute Reliabilität, ICC zwischen 0,88 und 0,92 für Messungen im Abstand einer Woche [Ostir 2002]. Gute Änderungssensitivität mit ES zwischen 0,48 und 1,25.			
Begründung der Aufnahme in die Leitlinie:			
Vielfältiger Prognosefaktor. Spielt eine Rolle in der Sarkopenie-Diagnostik. In Studien international breit vertreten.			
Limitationen / Forschungsbedarf:			
Bei stationär behandelten geriatrischen Patient*innen besteht ein deutlicher Bodeneffekt, bei sehr fitten älteren Menschen besteht ein Deckeneffekt [Kameniar 2022]. Bezüglich der Sturzprädiktion sind die Erkenntnisse inkongruent.			
Literatur:			
Beck Jepsen D, Robinson K, Ogliari G et al. (2022) Predicting falls in older adults: an umbrella review of instruments assessing gait, balance, and functional mobility. <u>BMC geriatrics</u> 22(1):615			
Bruyere O, Beaudarta C, Reginster JY et al. (2016) Assessment of muscle mass, muscle strength and physical performance in clinical practice: An international survey. <u>Eur Geriatr Med</u> 7:243-246			
Cruz-Jentoft AJ, Bahat G, Bauer J et al. (2019) Sarcopenia: revised European consensus on definition and diagnosis. <u>Age Ageing</u> , 48:16-31. DOI:10.1093/ageing/afy169			
Freiberger E, de Vreede P, Schoene D et al. (2012) Performance-based physical function in older community-dwelling persons: a systematic review of instruments. <u>Age Ageing</u> 41:712-721 DOI:			

10.1093/ageing/afs099

Guralnik JM, Simonsick EM, Ferrucci L et al. (1994) A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 49:M85-94.

Guralnik J, Ferrucci L, Pieper CF et al. (2000) Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 55:M221-M231

Kameniar K, Mackintosh S, Van Kessel G, Kumar S (2022) The Psychometric Properties of the Short Physical Performance Battery to Assess Physical Performance in Older Adults: A Systematic Review. J Geriatr Phys Ther. DOI: 10.1519/JPT.0000000000000337

Ostir GV, Volpato S, Fried LP et al. (2002) Reliability and sensitivity to change assessed for a summary measure of lower body function: results from the Women's Health and Aging Study. J Clin Epidemiol 55:916-921

Volpato S, Cavalieri M, Sioulis F et al. (2011) Predictive Value of the Short Physical Performance Battery Following Hospitalization in Older Patients. J Gerontol A Biol Sci Med Sci 66(1):89-96. DOI: 10.1093/gerona/glq167

Tab. 2.3.9: Dimension Mobilität

Name des Instrumentes: Kürzel:	Performance-Oriented Mobility Assessment nach Tinetti POMA (I, II)	
Mittlerer Zeitaufwand: Für Patient*in 15 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 17 min		
Lernaufwand für Untersucher*in:		
O leicht	O mittel	x anspruchsvoll
Durchführung durch: Physiotherapeut*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
x leicht/Früherkennung	x mittel	O schwer
Gütekriterien: Variables Alter in den verschiedenen Studienpopulationen (45-85 J), häusliche Umgebung oder Rehaklinik, meist nur schwache Korrelationen mit anderen Mobilitätstests, meist gute Interrater- und Retest-Reliabilität, schlechte prädiktive Wertigkeit für Stürze in einem systemischen Review [Omaña 2021].		
Begründung der Aufnahme in die Leitlinie: Von der AGAST empfohlenes Instrument, das erfahrenen Untersuchenden einen Überblick über unterschiedliche Parameter der Mobilität verschafft.		
Limitationen / Forschungsbedarf:		

Einsatz nur sinnvoll, wenn Person aufstehen kann. Anwendung muss mehrstündig trainiert werden. Nicht international standardisiert - verschiedene Testvarianten und Cut-off-Werte erschweren die Beurteilung der Gütekriterien. Es liegen meist nur kleine Studien zu spezifischen Krankheitsbildern vor (u. a. Schlaganfall, Parkinson, Huntington, Arthrose).

Literatur:

Canbeck J, Fulk G, Nof L, Echternach J (2013) Test-retest reliability and construct validity of the tinetti performance-oriented mobility assessment in people with stroke. J Neurol Phys Ther 37:14-19

Hayati M, Furtado GE, Nazarali P et al. (2025) Cross-sectional assessment of the Tinetti performance-oriented mobility tool for screening physical frailty syndrome in older adults. BMC Geriatr 25(1):214. DOI: 10.1186/s12877-025-05858-0

Köpke S, Meyer G (2006) The Tinetti test. Babylon in geriatric assessment. Z Gerontol Geriat 39:288-291

Omaña H, Bezaire K, Brady K et al. (2021) Functional Reach Test, Single-Leg Stance Test, and Tinetti Performance-Oriented Mobility Assessment for the Prediction of Falls in Older Adults: A Systematic Review. Phys Ther 101:pzab173. DOI: 10.1093/ptj/pzab173

Park J, Koh SB, Kim HJ et al. (2018) Validity and Reliability Study of the Korean Tinetti Mobility Test for Parkinson's Disease. J Mob Disord 11:24-29

Parveen H, Noohu MM (2017) Evaluation of psychometric properties of Tinetti performance-oriented mobility assessment scale in subjects with knee osteoarthritis. Hong Kong Physiotherapy Journal 36: 25-32

Schüle S (2014) Performance-Oriented Mobility Assessment und Berg Balance Scale im Vergleich. Z Gerontol Geriat 47:153-164

Schüle S, Pflugrad L, Petersen H et al. (2017) Deutsche Übersetzung des "Performance-Oriented Mobility Assessment" nach Tinetti. Z Gerontol Geriat 50:498-505

Tinetti M (1986) Performance-oriented assessment of mobility problems in the elderly. J Am Geriatr Soc 34:119

Tinetti M, Ginter S (1988) Identifying mobility dysfunction in the elderly. JAMA 259:1190-1193

Tab. 2.3.10: Dimension Mobilität

Name des Instrumentes:	Berg Balance Scale		
Kürzel:	BBS		
Mittlerer Zeitaufwand:			
Für Patient*in 24min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 26 min			
Lernaufwand für Untersucher*in:			
☐ leicht	☒ mittel	☐ anspruchsvoll	
Durchführung durch:			
Physiotherapeut*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☒ schwer	

Gütekriterien:

Eine deutsche Übersetzung ist verfügbar [Scherfer 2006].

Beck Jepsen (2022): Umbrella Review über 31 Reviews zum Assessment der Mobilität. In 3 Reviews wurde der Einsatz der BBS eher befürwortet, am ehesten nach Insult, in 4 wurde der Einsatz eher kritisch gesehen.

Lusardi (2017): Systematischer Review zum Sturzrisiko. Von 119 Tests wies die BBS beim Cut-off 50/49 den höchsten positiven Vorhersagewert auf.

und Normwerte [Downs 2014]. Gute prädiktive Validität für den Beginn von Einschränkungen der Alltagsaktivitäten [Downs 2014]. In einem älteren Review konnte die Validität und Reliabilität der BBS bei Schlaganfallpatient*innen nachgewiesen werden, es zeigten sich aber auch Boden- und Deckeneffekte [Blum 2008].

Die Retest-Reliabilität lag bei 0,95 (95% Konfidenzintervall 0,85–0,98) und die anhand gepoolter Daten geschätzte Intrarater-Reliabilität war hoch (0,98; 95% Konfidenzintervall 0,97–0,99). Bei Untersuchung von Patient*innen mit Alzheimer-Demenz betrug die Interrater-Reliabilität 0,72 (95% Konfidenzintervall 0,31–0,91) und die Retest-Reliabilität wurde mit 0,95 (95% Konfidenzintervall 0,85–0,98) angegeben [Muir 2008].

Begründung der Aufnahme in die Leitlinie:

Erfahrenen Untersuchenden ermöglichen die 14 Items eine differenzierte Beurteilung der Mobilität hinsichtlich ihrer auf das Gleichgewicht bezogenen Aspekte. Erfassung mittels tragbarer Sensoren möglich [Lu 2025].

Limitationen / Forschungsbedarf:

Nicht bei Demenz oder geistiger Behinderung einsetzbar. Moderate Veränderungen zeigt die BBS nicht zuverlässig an [Downs 2013]. Auch ein normaler Score schließt ein erhöhtes Sturzrisiko nicht aus [Muir 2008].

Literatur:

Beck Jepsen D, Robinson K, Ogliari G et al. (2022) Predicting falls in older adults: an umbrella review of instruments assessing gait, balance, and functional mobility. BMC Geriatrics 22(1):615 <https://doi.org/10.1186/s12877-022-03271-5>

Berg K, Wood-Dauphinee S, Williams JI, Gayton D (1989) Measuring balance in the elderly: preliminary development of an instrument. Physiother Can 41:304-311 DOI: 10.3138/ptc.41.6.304

Blum L, Korner-Bitensky N (2008) Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. Phys Ther 88:559-566

Dolatbadi E, Van Ooteghem K, Taati B, Iaboni A (2018) Quantitative mobility assessment for fall risk prediction in dementia: a systematic review. Dement Geriatr Cogn Disord 45(5–6):353-367

Downs S (2015) The Berg Balance Scale. J Physiother 61:46

Downs S (2014) Normative scores on the Berg Balance Scale decline after age 70 years in healthy community-dwelling people: a systematic review. J Physiother 60:85-89

Downs S, Marquez J, Chiarelli P (2013) The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review. J Physiother 59:93-99

Lee J, Geller AI, Strasser DC (2013) Analytical review: focus on fall screening assessments. PM R. 5(7):609-621. DOI: 10.1016/j.pmrj.2013.04.001

Lima CA, Ricci NA, Nogueira EC, Perracini MR (2018) The Berg Balance Scale as a clinical screening tool to predict fall risk in older adults: a systematic review. Physiotherapy 104(4):383-394

Lu Z, Zhou H, Lyu H et al. (2025) Berg Balance Scale scoring system for balance evaluation by leveraging attention-based deep learning with wearable IMU sensors. Bioengineering 12:395. <https://doi.org/10.3390/bioengineering12040395>.

Lusardi MM, Fritz S, Middleton A et al. (2017) Determining Risk of Falls in Community Dwelling Older Adults: A Systematic Review and Meta-analysis Using Posttest Probability. J Geriatr Phys Ther 40(1):1-36. DOI: 10.1519/JPT.0000000000000099

Muir SW, Berg K, Chesworth B, Speechley M (2008) Use of the Berg Balance Scale for predicting multiple falls in community-dwelling. Phys Ther 88:449–459

Neuls PD, Clark TL, Van Heuklon NC et al. (2011) Usefulness of the Berg Balance Scale to predict falls in the elderly. J Geriatr Phys Ther 34(1):3-10

Park SH (2018) Tools for assessing fall risk in the elderly: a systematic review and meta-analysis. Aging Clin Exp Res 30(1):1-16. DOI: 10.1007/s40520-017-0749-0

Scherfer E, Bohls C, Freiburger E et al. (2006) Berg-Balance-Scale – deutsche Version. Übersetzung eines Instruments zur Beurteilung von Gleichgewicht und Sturzgefährdung. Physiotherapie 2:1-8

Tab. 2.3.11: Dimension Mobilität

Name des Instrumentes:	Falls Efficacy Scale-International und Icon-Version		
Kürzel:	FES-I / Icon-FES		
Mittlerer Zeitaufwand:			
Für Patient*in 11 min, Icon-FES 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 13min bzw. 12 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Psycholog*in, Ärzt*in, Physiotherapeut*in, Ergotherapeut*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien:			
Validierte Übersetzung. Gute prädiktive Validität für Stürze bei hinsichtlich Alter, Geschlecht und Sturzbiographie unterschiedlichen Gruppen.			
McGarrigle (2023) (systematisches Review): gute psychometrische Werte bei Verwendung bei einer großen Bandbreite älterer Personen (gesunde älteren Menschen, Menschen mit einem hohen Sturzrisiko, z. B. mit Erkrankungen, die die Mobilität und das Gleichgewicht beeinträchtigen): Gute bis sehr gute Retest-Reliabilität (ICC 0,79–0,99); gute bis sehr gute Interrater-Reliabilität (ICC 0,72–0,98); gute bis sehr gute interne Konsistenz (Cronbachs α : 0,78–0,99).			
Hauer (2011): Änderungssensitivität: Standardized Response Mean (SRM) 0,21–0,74 (Spannweite), Bodeneffekt bei kognitiver Einschränkung (MMSE<24) 3,2%, ohne kognitive Einschränkung 4,7%, keine Deckeneffekte.			
Delbaere (2010): Cut-off geringe/hohe Bedenken bzw. Angst zu stürzen 16-22/23-64 Punkte, geringe/moderate/hohe Bedenken/Angst zu stürzen 16-19/20-27/28-64 Punkte.			

Ikongrafische Version:

Delbaere (2011): $n = 250$, $80,2 \pm 5,1$ (70-90) J, 53,2% F, ambulant, mindestens 20 m gehfähig, ohne Demenz. Interne Konsistenz Cronbachs α 0,96, Korrelation mit FES-I hinsichtlich des Summenscores Spearmans rho 0,742, auf Item-Ebene 0,411 bis 0,617, Bodeneffekt 4%, Retest-Reliabilität ($n = 50$, Intervall 1 Woche) ICC 0,90.

Begründung der Aufnahme in die Leitlinie:

Sturzbedenken und Sturz-assoziierte Selbstwirksamkeit können unabhängig von der physischen Funktion zu einem ausgeprägten Vermeidungsverhalten und damit Immobilität führen. Der FES-I ist das am besten validierte Erfassungsinstrument.

Limitationen / Forschungsbedarf:

Sensitivität und Spezifität sind mangels Vergleichsstandard für Sturzbedenken bislang nicht erhoben worden.

Zukünftige Forschung muss noch die Änderungssensitivität des Fragebogens erheben. Werte hierzu gibt es aktuell noch kaum. Auch der Wechsel von einem „schützenden“ Bedenkenlevel hin zu einem „maladaptiven“ Bedenkenlevel ist noch nicht geklärt. „Schützende“ Bedenken können Stürze verhindern, während „maladaptive“ Bedenken das Sturzrisiko erhöhen [Ellmers 2023]. Für das Krankenhaus-Setting wird gerade eine entsprechende Version des FES-I entwickelt, die dort relevante Aktivitäten beinhaltet.

Literatur:

<https://sites.manchester.ac.uk/fes-i/>

https://www.neura.edu.au/wp-content/uploads/2016/05/Icon-FES_30item-house.pdf

Delbaere K, Close JC, Mikolaizak AS et al. (2010) The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. *Age Ageing* 39:210-216

Delbaere K, Close JC, Taylor M et al. (2013) Validation of the Iconographical Falls Efficacy Scale in cognitively impaired older people. *The journals of gerontology Series A, Biological sciences and medical sciences* 68(9):1098-1102

Delbaere K, Smith ST, Lord SR (2011) Development and Initial Validation of the Iconographical Falls Efficacy Scale. *J Gerontol A Biol Sci Med Sci* 66A:674-680

Denkinger MD, Igl W, Coll-Planas L et al. (2009) Practicality, validity and sensitivity to change of fear of falling self-report in hospitalised elderly—a comparison of four instruments, *Age and Ageing* 38(1):108–112. <https://doi.org/10.1093/ageing/afn23>

Dias N, Kempen GI, Todd CJ et al. (2006) The German version of the Falls Efficacy Scale-International Version (FES-I). *Z Gerontol Geriatr* 39:297-300

Ellmers TJ, Freiburger E, Hauer K et al. (2023). Why should clinical practitioners ask about their patients' concerns about falling? *Age Ageing* 52(4). DOI 10.1093/ageing/afad057

McGarrigle L, Yang Y, Lasrado R et al. (2023) A systematic review and meta-analysis of the measurement properties of concerns-about-falling instruments in older people and people at increased risk of falls. *Age Ageing* 52:1-11

Hauer KA, Kempen GI, Schwenk M et al. (2011) Validity and sensitivity to change of the falls efficacy scales international to assess fear of falling in older adults with and without cognitive impairment. *Gerontology* 57:462-472

Kempen GI, Todd CJ, Van Haastregt JC et al. (2007) Cross-cultural validation of the Falls Efficacy Scale International (FES-I) in older people: results from Germany, the Netherlands and the UK were satisfactory. *Disabil Rehabil* 29:155-162

Lim ML, van Schooten KS, Radford KA et al. (2021) The Iconographical Falls Efficacy Scale (IconFES) in community-dwelling older people—a longitudinal validation study. *Age Ageing*

50(3):822-829

Şentürk AY, Livanelioğlu A (2022) The Cross-Cultural Adaptation and Factor Analysis of the Turkish Version of Iconographical Falls Efficacy Scale Long Version. International Journal of Older People Nursing, e12453. <https://doi.org/10.1111/opn.12453>

Yardley L, Beyer N, Hauer K et al. (2005) Development and initial validation of the Falls Efficacy Scale-International (FES-I). Age Ageing 34:614-619

Tab. 2.3.12: Dimension Mobilität

Name des Instrumentes: Kürzel:	Falls Efficacy Scale-International kurz und Icon-Version Short FES-I / Short Icon-FES	
Mittlerer Zeitaufwand: Für Patient*in 5 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 7 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch: Psycholog*in/, Ärzt*in, Physiotherapeut*in, Ergotherapeut*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien: Hohe prädiktive Validität für Stürze bei denselben Gruppen wie FES-I, aber etwas geringer. Nach Delbaere et al. Cut-off geringe/hohe Bedenken bzw. Angst zu stürzen 1-10/11-28 Punkte, geringe/moderate/hohe Bedenken/Angst zu stürzen 7-8/9-13/14-28 Punkte. Retest-Reliabilität bei Durchführung im Abstand von 4 Wochen über ICC 0,83. Änderungssensitivität nicht so gut wie in der Langversion, Standardized Response Mean 0,18–0,91 (Spannweite). Interne Konsistenz über Cronbachs α nach McGarrigle 0,63 - 0,98. Bodeneffekt bei kognitiver Einschränkung (MMSE<24) 14%, ohne kognitive Einschränkung 11% [Hauer 2011]. Keine Deckeneffekte.		
<u>Ikongrafische Version:</u> Delbaere (2011): (Studienpopulation siehe Icon-FES der Langversion) ähnliche Gütekriterien, interne Konsistenz Cronbachs α 0,87, hohe Retest-Reliabilität mit ICC 0,92–0,93.		
Begründung der Aufnahme in die Leitlinie: Die Kurzform des FES-I ist aufgrund der Praktikabilität das mit großem Abstand am meisten genutzte Instrument bei nur bei leichtem Abfall der Güteeigenschaften, ansonsten siehe FES-I.		

Limitationen / Forschungsbedarf:

Bei höhergradiger Demenz nicht einsetzbar. Bestimmte anspruchsvolle Aktivitäten in Bezug auf die Abfrage von Angst fehlen, wenn es um die Planung für Therapien geht.

Literatur:

<https://sites.manchester.ac.uk/fes-i/>

Delbaere K, Close JC, Mikolaizak AS et al. (2010) The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. Age Ageing 39:210-216

Delbaere K, Smith ST, Lord SR (2011) Development and Initial Validation of the Iconographical Falls Efficacy Scale. J Gerontol A Biol Sci Med Sci 66A:674-680

Delbaere K, Smith ST, Lord SR (2013) Iconographical Falls Efficacy Scale. American Psychological Association. <https://doi.org/10.1037/t15733-000>

Hauer KA, Kempen GI, Schwenk M et al. (2011) Validity and sensitivity to change of the falls efficacy scales international to assess fear of falling in older adults with and without cognitive impairment. Gerontology 57:462-472

Kempen GI, Yardley L, van Haastregt JC et al. (2008) The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age Ageing 37:45-50

Lim ML, van Schooten KS, Radford KA et al. (2021) The Iconographical Falls Efficacy Scale (IconFES) in community-dwelling older people—a longitudinal validation study. Age Ageing 50(3):822-829

McGarrigle L, Yang Y, Lasrado R et al. (2023) A systematic review and meta-analysis of the measurement properties of concerns-about-falling instruments in older people and people at increased risk of falls. Age Ageing 52:1-11

Beispiele:

1. 90-jährige Patient*in, anamnestisch und klinisch ohne Hinweis auf beeinträchtigte Mobilität,
TUG <15 Sekunden
→ **derzeit keine Indikation zu weiterem Assessment der Mobilität**
2. Patient*in mit atropher Oberarm- und Wadenmuskulatur,
TUG 30 Sekunden am Rollator
→ **Handkraftmessung und 5-Chair-Rise Test**
3. Stationäre Patient*in, gehfähig, aber kann TUG nicht bewältigen, da Aufstehen nur mit personeller Hilfe möglich ist
→ **4 m Gehtest**
4. Patient*in nicht ohne personelle Hilfe gehfähig
→ **LSBM**
5. Patient*in ist mehrfach gestürzt (→somatische Diagnostik zur Ursachenklärung),
TUG 12 Sekunden ohne Hilfsmittel
→ **Balance-Tests frei kombiniert (z. B. Romberg-Stand, Einbeinstand usw.),
TUG dual task (erhöhte Sturzgefahr unter Ablenkung?),
ggf. FES-I short/long (ausgeprägte Sturzangst?)**

Tab. 2.3.13: Motorik (obere Extremität)

Name des Instrumentes:	8-Punkte-Greifraum-Test		
Kürzel:	8P-GRT		
Mittlerer Zeitaufwand: Für Patient*in 2 min (1 pro Seite) / für Untersucher*in (inkl. Vor- u. Nachbereitung) 3 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch: Ärzt*in, Ergotherapeut*in, Physiotherapeut*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☐ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien: <u>Krupp (2022)</u> : Studienpopulation n = 82, 81,6 ± 6,4 (64 – 96) J, 62% F, stationäre Akutgeriatrie. Interrater-Reliabilität rechte Seite 0,99 (0,99 – 1,00), linke Seite 0,99 (0,98 – 1,00), Retest-Reliabilität rechte Seite 0,95 (0,91 – 0,98), linke Seite 0,90 (0,81 – 0,96), kein signifikanter Einfluss von Alter, Geschlecht und Händigkeit. Interne Konsistenz Cronbachs α rechts 0,78, links 0,76. Korrelation der Werte der rechten Seite mit Fragebogen zu korrespondierenden Basis-Aktivitäten der Selbstversorgung -0,72, Shoulder Pain and Disability Index (Anteil Disability) -0,60, 20-Cents-Test 0,40, Handkraft 0,39. Deckeneffekt (= beidseits 8) bei 4,1% (n = 3), kein Bodeneffekt. <u>Krupp (2019)</u> : Studienpopulation n=255, 82,8 ± 8,7 (52 – 99) J, 77,7% F, Pflegeheimbewohner*innen. Feasibility auch bei mittelschwerer Demenz (Six-Item Screener 0 Punkte) gegeben.			
Begründung der Aufnahme in die Leitlinie: Performance-Test, der seitengetrennt erfasst, ob die Positionierung der Hand alltagsrelevante Aktivitäten ermöglicht, auch bei höhergradiger kognitiver Störung einsetzbar.			
Limitationen / Forschungsbedarf: Da geeignete Performance-Tests nicht als Goldstandard zur Verfügung stehen und geriatrische Patient*innen defezitäre Selbstversorgungsfähigkeiten häufig nicht kausal zuordnen können, ist eine Berechnung der Sensitivität/Spezifität nicht möglich.			
Literatur: Deutsche Version in: <u>Krupp S, Petersen T, Balck F et al. (2022)</u> Die Hand positionieren, um zu agieren – Validierung des 8-Punkte-Greifraum-Tests. Z Gerontol Geriat 55:93–98. https://doi.org/10.1007/s00391-022-02029-3 Dissertation Petersen T (2021) https://www.zhb.uni-luebeck.de/epubs/ediss2893.pdf <u>Krupp S, Kasper J, Hermes A. et al. (2019)</u> Das „Lübecker Modell Bewegungswelten“ – Ergebnisse der Effektevaluation. Bundesgesundheitsbl 62:274–281. https://doi.org/10.1007/s00103-019-02881-6			

Tab. 2.3.14: Motorik (obere Extremität)

Name des Instrumentes:	Handkraftmessung (Greifkraft)		
Kürzel:	–		
Mittlerer Zeitaufwand: Für Patient*in 6 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 8 min			
Lernaufwand für Untersucher*in:			
x leicht		O mittel	O anspruchsvoll
Durchführung durch: Physiotherapeut*in, Ergotherapeut*in, Ärzt*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
x leicht/Früherkennung		x mittel	x schwer
Gütekriterien: Daten verschiedener Bevölkerungsgruppen erlauben einen Vergleich der ermittelten Patientenwerte mit der Norm [Huemer 2023, dabei älteste Referenzgruppe 70-74 J; Steiber 2016]. Gagesch (2023): n = 976, 75,2 ± 4,6 J, 62% F, gesunde Schweizer*innen (DO-HEALTH Studie). Messung mittels Vigorimeter. Vier verschiedene statistische Methoden der Definition klinisch relevanter muskulärer Schwäche wurden für die Ableitung des Cut-offs nach Geschlecht und Alterskategorie (≤75 vs. >75 Jahre) verwendet. Benton (2022): n = 67, 76 ± 1 J, 49% F, nicht-institutionalisiert wohnende, gehfähige Personen. Vergleich der Werte gemäß Dynamometer nach Jamar und Smedley. Jeweils ein Versuch. Signifikant höhere Werte mit der Methode nach Jamar, insbesondere bei Männern und Teilnehmenden bis 75 J. Retest-Reliabilität über ICC 0,98 (Jamar) bzw. 0,96 (Smedley). Bobos (2020): Hohe Retest-Reliabilität über ICC 0,92 bei gesunden Teilnehmenden, 0,95 nach Verletzungen der oberen Extremität, 0,96 bei neurologisch Erkrankten. Kleinste klinisch relevante Änderung bei Schlaganfallpatienten 5,0 kg (dominante Seite) und 6,2 kg (nichtdominante Seite), 6,5 kg nach distaler Radiusfraktur und 2,44-2,69 kg bei Gesunden.			
Begründung der Aufnahme in die Leitlinie: Alltagsrelevanter Aspekt der Handfunktion und vielfältiger Prognosefaktor. Spielt eine Rolle in der Sarkopenie-Diagnostik.			
Limitationen / Forschungsbedarf: Für Verlaufsmessungen sollte ein weitgehend identisches Instrument verwendet werden (Dynamometer oder Vigorimeter). Relativ hohe intraindividuelle Streubreite der Messwerte.			
Literatur: Benton MJ, Spicher JM, Silva-Smith AL (2022) Validity and reliability of handgrip dynamometry in older adults: A comparison of two widely used dynamometers. PLoS ONE 17(6):e0270132. https://doi.org/10.1371/journal.pone.0270132 Beseler MR, Rubio C, Duarte E et al. (2014) Clinical effectiveness of grip strength in predicting			

ambulation of elderly inpatients. Clin Interv Aging 9:1873-1877

Bobos P, Nazari G, Lu Z, MacDermid JC. (2020) Measurement Properties of the Hand Grip Strength Assessment: A Systematic Review With Meta-analysis. Arch Phys Med Rehabil101(3):553-565

Bohannon RW (2019) Grip Strength: An indispensable biomarker for older adults. Clin Interv Aging 14:1681–1691

Bohannon RW (2015) Muscle strength: clinical and prognostic value of hand-grip dynamometry. Curr Opin Clin Nutr Metab Care 18:465-470

Gagesch M, Wieczorek M, Abderhalden LA et al. (2023) Grip strength cut-points from the Swiss DO-HEALTH population. Eur Rev Aging Phys Act 20(1):13. DOI:10.1186/s11556-023-00323-6

Huemer MT, Kluttig A, Fischer B et al. (2023) Gripstrength values and cut-off pointsbased on over 200,000 adults of the German National Cohort - a comparison to the EWGSOP2 cut-off points. Age Ageing 52:1-10. DOI: 10.1093/ageing/afac324

Jeong W, Moon JY, Kim JH (2023) Association of absolute and relative hand grip strength with all-cause mortality among middle-aged and old-aged people. BMC Geriatr 23:321

Laukkanen JA, Voutilainen A, Kurl S et al. (2020) Handgrip strength is inversely associated with fatal cardiovascular and all-cause mortality events. Ann Med 52(3-4):109-119. DOI: 10.1080/07853890.2020.1748220

Mayhew AJ, So HY, Ma J et al. (2023) Normative values for grip strength, gait speed, timed up and go, single leg balance, and chair rise derived from the Canadian longitudinal study on ageing. Age Ageing 52(4). DOI:10.1093/ageing/afad054

Philipps P (1986) Grip strength, mental performance and nutritional status as indicators of mortality risk among female geriatric patients. Age Ageing 15:53-56

Reuter SE, Massy-Westropp N, Evans AM (2011) Reliability and validity of indices of hand-grip strength and endurance. Aust Occup Ther J 58:82-87

Roberts HC, Denison HJ, Martin HJ et al. (2011) A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age Ageing 40:423-429

Steiber N (2016) Strong or weak handgrip? Normative reference values for the German population across the life course stratified by sex, age, and body height. PLoS One 4;11(10):e0163917

Tab. 2.3.15: Motorik (obere Extremität)

Name des Instrumentes:	20-Cents-Test (20-Coins-Test)	
Kürzel:	20-C-T	
Mittlerer Zeitaufwand:		
Für Patient*in 2 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Ärzt*in, Ergotherapeut*in, Physiotherapeut*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☒ schwer

Gütekriterien:

Krupp (2022): In der Validierung eines Fragebogens zur Selbstbeurteilung von bei Alltagsaktivitäten auftretenden Schwierigkeiten mit Bezug zur Handfunktion (Duruöz Hand-Index) bei 101 tagesklinisch oder ambulant behandelten geriatrischen Patient*innen (80 ± 6 J, 73% F) war die Korrelation des Index mit dem 20-C-T beidseits höher (rechts -0,274, $p = 0,009$; links -0,286, $p = 0,005$) als die Korrelation mit der Handkraft (rechts -0,209, $p = 0,049$; links mit -0,174, $p = 0,102$ nicht signifikant).

Krupp (2015): Studienpopulation $n = 300$, 62,4% stationäre Akutgeriatrie, Rest Tagesklinik und ambulante Rehabilitation; $n = 205$ für die Validität, $79,2 \pm 7,4$ (60 – 96) J, 65,4% F; $n = 95$ für die Reliabilität, $80,6 \pm 7,3$ (61 – 95) J, 69,4% F. Cut-off je nach Intention der Testung (Ausschluss versus Nachweis einer Störung). Sensitivität für das Erkennen einer Störung der Feinmotorik, die im Alltag nicht voll kompensiert werden kann, bei Cut-off von 20 Cents in 40/>40 Sekunden 87%, Spezifität 77%. Kein signifikanter Einfluss von Händigkeit, Geschlecht, Alter, geringer Einfluss der Kognition (partiell $\eta^2 = 0,049$ bei MMSE 12-30). Intrarater-Reliabilität: Cronbachs α 0,956 (95%-Konfidenzintervall 0,942–0,967). Interrater-Reliabilität: Cronbachs α 0,933 (95%-Konfidenzintervall 0,918–0,946).

Begründung der Aufnahme in die Leitlinie:

Performance-Test mit seitengetrennter Beurteilung alltagsrelevanter feinmotorischer Aspekte der Handfunktion, auch bei mittelgradiger kognitiver Störung einsetzbar.

Limitationen / Forschungsbedarf:

Bei schwerer Demenz nicht zuverlässig. Bodeneffekt (0 Cents in 60 Sekunden transferiert) auf die einzelne Hand bezogen 2% (8 von 407 untersuchten Händen), bei keiner untersuchten Person beidseits. Forschungsbedarf in Bezug auf die Änderungssensitivität. (Im nahezu identischen 10 Coins Test (Yamada 2025) stetige Abnahme des Zeitbedarfs im ersten Halbjahr nach Operation bei zervikaler spondylotischer Myelopathie.)

Literatur:

Deutsche Version in:

Krupp S, Kasper J, Balck F et al. (2015) „Timed up and go“ für die Finger in Form des 20-Cents-Tests. Psychometrische Gütekriterien eines einfachen Feinmotorik-Performance-Tests. Z Gerontol Geriat 48:121–127. <https://link.springer.com/article/10.1007/s00391-014-0854-z>

Krupp S, Schöne F, Balck F et al. (2022) Selbstbeurteilung von Alltagsaktivitäten mittels Duruöz-Hand-Index – Validierung der deutschen Übersetzung Z Gerontol Geriat 55(2):99-104. DOI: 10.1007/s00391-022-02041-7

Yamada S, Izumi T, Inokuchi A et al. (2025) New assessment method for upper limb function in patients with cervical spondylotic myelopathy. The 10 Coins Test. Spine 50(3):145-155. DOI: 10.1097/BRS.0000000000005179

2.4 Kognition / Delirerfassung

Kognition

Der **Six-Item Screener (SIS)** (Tabelle 2.4.1) als einminütiger Test der Stufe 2a ohne Materialeinsatz eignet sich besonders für die Aufnahmesituation, da er nur minimal belastet und auch bei schwerer feinmotorischer oder Sehstörung sowie unter intensivierten Hygieneregeln unkompliziert einsetzbar ist. So fällt bereits vor der ersten

Nacht eine das Delirrisiko steigernde alltagsrelevante kognitive Störung auf. Die Kenntnis des Ergebnisses des SIS erleichtert es, bei unter 6 Punkten oder klinischem/anamnestischem Hinweis auf das Vorliegen eines kognitiven Defizits zu passender Zeit ein weiteres Assessment-Instrument sinnvoll zu kombinieren: Oberhalb von 3 Punkten machen zur Früherkennung geeignete Verfahren (auch Uhrentests) mehr Sinn als die Mini Mental State Examination (MMSE). Unterhalb von 2 Punkten reicht die Kooperationsfähigkeit der Untersuchten oft nicht für die Erfassung aller Items der MMSE, so dass die Severe Mini Mental State Examination (SMMSE) mehr Vorteile bieten kann.

Das Ergebnis von **Uhr-Zeichen-Tests (Clock Drawing Test, CDT)** (Tabellen 2.4.2, 2.4.3 und 2.4.4) hängt von verschiedenen kognitiven Fähigkeiten ab: selektiver und anhaltender Aufmerksamkeit, auditivem Verständnis, verbalem Arbeitsgedächtnis, Fähigkeit im Umgang mit Zahlen, räumlichem Gedächtnis, visuokonstruktiven Fähigkeiten und exekutiven Funktionen [Mendez 1992]. Es existieren mehr als 30 verschiedene qualitative, semi-quantitative und quantitative Auswertungssysteme, über deren Wertigkeit bis jetzt kein Konsens besteht [Spenciare 2017]. Erhöhte Komplexität der Auswertung erbringt nach bisherigem Wissen keinen wesentlichen Vorteil in Bezug auf die Sensibilität in der Demenzerkennung [Mainland 2014]. Manche Systeme beziehen die gesamte Uhr in die Auswertung ein (vor allem qualitative) und manche nur Teilaspekte. Insofern unterscheiden sie sich in der Erkennung verschiedener Schweregrade kognitiver Beeinträchtigungen. Für alle Uhr-Zeichen-Tests gilt, dass die Beobachtung der Performance wichtige Zusatzinformationen für die Interpretation des grafischen Ergebnisses liefert.

Der **Uhr-Zeichen-Test nach Shulman** (Tabelle 2.4.2) wurde ursprünglich zur Beschreibung neuropsychologischer Defizite, speziell der räumlich konstruktiven Fähigkeiten genutzt. Shulman et al. veröffentlichten erstmals 1986 eine Studie, die den CDT als Screening-Tool für Demenz bei älteren Menschen untersuchte und modifizierte die Bewertung 1993 (Score 1 bis 6). Der Test ist rasch durchführbar, einfach anzuwenden, gut akzeptiert von Patient*innen, einfach auszuwerten und relativ unabhängig von sprachlichen Fähigkeiten, Schulbildung und kulturellem Hintergrund. Obwohl es sich um die einfachste und älteste publizierte Auswertungsmethode handelt, zeigt der Uhr-Zeichen-Test nach Shulman im Vergleich zu anderen Auswertungsmethoden sehr hohe Werte für Sensitivität und Spezifität in der Demenzerkennung. Dies kann durch Kombination mit einem weiteren Demenztest (z.B. MMSE) noch gesteigert werden. Eine Differenzierung verschiedener Demenzformen anhand des CDT wurde vielfach versucht, erbrachte jedoch keine praktisch verwertbaren Ergebnisse. Lediglich fronto-temporale Demenzformen scheinen in quantitativen Analysen höhere Punktwerte aufzuweisen, was durch eine frühzeitige Störung exekutiver Funktionen erklärbar scheint [Tan 2015].

Der **Uhr-Zeichen-Test nach Sunderland** (Tabelle 2.4.3) ist bei frei gezeichnetem Kreis und der Zeitvorgabe 11:10 Uhr (zehn Minuten nach elf) ein sehr gut geeignetes, einfach und schnell anwendbares Verfahren. Der Test ist frei von kulturellen und bildungsabhängigen Einflüssen, nicht Angst erzeugend und wird von den Patient*innen meist sehr gut akzeptiert und toleriert. Die Auswertung nach Sunderland mit Vergabe von 1 bis 10 Punkten ist im TFDD (s. u.) enthalten und bewährte sich bei mehr als 13500 neuropsychologischen Untersuchungen im Gerontopsychiatrischen Zentrum des PSD

Wien sehr gut.

Das **Uhren Ergänzen (Clock Completion, CC) nach Watson** (Tabelle 2.4.4) wurde historisch nach den Erstveröffentlichungen der Verfahren von Shulman (1986) und Sunderland (1989) im Jahr 1993 publiziert und von AGAST 1995 für das Basisassessment empfohlen [AGAST 1995]. In einen vorgegebenen Kreis (Durchmesser: 10 cm) werden die Ziffern einer Uhr eingezeichnet. Eine zeitliche Vorgabe für die Ausführung besteht nicht. Die Auswertung erfolgt nach vier gleich verteilten Quadranten, die von der Ziffer 12 ausgehend z. B. als Folie über die gezeichnete Uhr gelegt werden. Fehler in den Quadranten 1 bis 3 werden mit je einem Punkt bewertet, ein Fehler im letzten, 4. Quadranten gewichtet mit gleich 4 Fehlerpunkten (maximal mögliche Anzahl von 7 Fehlerpunkten). Bei mehr als 3 Fehlerpunkten wird eine kognitive Störung angenommen, z.B. ein Demenzsyndrom. Das Zeichnen von Zeigern geht nicht in den Bewertungsscore ein. Die Testgütekriterien sind den Verfahren von Shulman und Sunderland unterlegen. Es ist daher unverständlich und inakzeptabel, diesen Uhrentest alternativlos für das Assessment im Rahmen der Verordnung geriatrischer Rehabilitation in Deutschland zu fordern, wie im „Muster 61“ vorgesehen.

Historisch gehören Uhr-Zeichen-Tests zu den am häufigsten eingesetzten Papier-und-Bleistift-Scores überhaupt (erstmalig im Jahr 1915 beschrieben). Ein Überblick zählt über 2.000 Publikationen seit 1989 [Hazan 2018]. Zusammenfassend sagt die S3-Leitlinie Demenzen aber: „Die Sensitivität dieser Verfahren ist begrenzt und sie sind zur Differenzialdiagnostik nicht geeignet.“ [Jessen 2017, S. 37]. Die S3-Leitlinie Demenzen empfiehlt ausdrücklich, bei der Einschätzung kognitiver Störungen Uhr-Zeichen-Tests nur gemeinsam mit z. B. MMSE, DemTect, TFFD oder MoCA durchzuführen – in beiden letztgenannten bereits enthalten (siehe oben). Aktuell wird diskutiert, ob die Aussagekraft von Uhr-Zeichen-Tests dadurch verbessert werden könnte, dass nicht nur das Ergebnis der fertig durchgeführten Aufgabe, sondern Exekutivfunktionen (z. B. auch Kompensationsstrategien) während der Ausführung im Ablauf beobachtet und spezifiziert werden [Spenciere 2017].

Der **Mini-Cog** (Tabelle 2.4.5) besteht aus dem Uhr-Zeichen-Test nach Sunderland und dem 3-Item Recall. Für eine fehlerlose Uhr inkl. korrekter Zeitwiedergabe von 11:10 Uhr werden 2 Punkte vergeben (die gespreizte Zeigerstellung scheint eine größere Herausforderung zu sein als andere Zeitangaben bzw. frei gewählte Zeitangaben) und für jedes korrekt erinnerte Wort 1 Punkt. Der Höchstwert beträgt also 5 Punkte, bei unter 3 Punkten besteht der Verdacht auf ein relevantes kognitives Defizit. Bei fremdsprachigen Patient*innen mit niedrigem Bildungsgrad bietet der Test den Vorteil einer deutlich besseren Performance im Vergleich zu z. B. der MMSE.

Der **„Schnelle Uhren-Dreier“ (SUD)** (Tabelle 2.4.6) ist eine Weiterentwicklung des Mini-Cog für den deutschen Sprachraum. Das Verfahren verbindet den Uhr-Zeichen-Test nach Sunderland mit dem 3-Item Recall und ist für ein rasches, erstes Demenzscreening sehr gut geeignet. Die Auswertung erfolgt mit einem klaren, anschaulichen Farbsystem (grün - gelb - rot).

Jedes der obig geschilderten Instrumente des kognitiven Assessments erreicht solitär lediglich die Tiefe des Assessments der Stufe 2a, d. h., es soll bereits bei leichter Auffälligkeit oder Vorliegen anamnestischer oder klinischer Hinweise auf das Vorliegen einer evtl. therapielevanten kognitiven Beeinträchtigung durch ein weiteres

Assessment-Verfahren ergänzt werden.

Der **Brief Alzheimer Screen (BAS)** (Tabelle 2.4.7) kombiniert zeitliche Orientierung (Datum), 3-item recall, Konzentration (backwards spelling) und Wortflüssigkeitsaufgabe (animal 30-score) zu einem sensitiven Test für die Detektion einer leichten Demenz. Der Test kommt ohne Lesen, Schreiben und Zeichnen aus und ist daher auch bei Patient*innen mit schwerer Sehbehinderung und/oder beeinträchtigter Feinmotorik geeignet. Der Test zeigt eine hohe Patientenakzeptanz und wird vom Memory Screening Advisory Board der Alzheimer's Foundation of America (AFA) als ein effektives, sensitives, spezifisches, leicht anwendbares und validiertes Instrument empfohlen.

Der **Bamberger Demenz-Screening Test (BDST)** (Tabelle 2.4.8) wurde mit dem Ziel entwickelt, auch von solchen geriatrischen Patient*innen gültige Ergebnisse zu erhalten, die z. B. aufgrund ihrer motorischen oder sensorischen Einschränkungen schlecht zeichnen oder schreiben können. Der BDST ist in wenigen Minuten als Bedside-Test ohne Formular durchführbar. Er enthält Aufgaben zur verbalen Flüssigkeit, zur kognitiven Flexibilität und zur konstruktiven Praxis und liefert unter anderem Informationen über die semantische, verbale und visuell-räumliche episodische Gedächtnisleistung. Die Autoren empfehlen ein adaptives Testen, bei dem Patient*innen mit eindeutigen kognitiven Defiziten schon bei den ersten Items nicht den gesamten Test zu absolvieren brauchen.

Die **Mini Mental State Examination (MMSE)** (Tabelle 2.4.9) ist das weltweit am häufigsten verwendete Verfahren der klinischen Demenzdiagnostik und wurde in mehr als 60 Sprachen übersetzt. Sie ist Teil umfangreicher Assessmentinstrumente, z. B. CERAD-NP Plus oder Addenbrooke's Cognitive Examination (ACE). Zur Verordnung von Antidementiva wird das Verfahren von vielen Kostenträgern gefordert und ist somit de facto unentbehrlich. Seit 2001 werden für die Verwendung des Tests Gebühren gefordert. Der MMSE-Test beinhaltet 30 Fragen. Zeitliche und örtliche Orientierung, Arbeitsgedächtnis und Aufmerksamkeit, Neugedächtnis (Einspeicherfähigkeit, verzögerter Abruf), Benennen, Lesen und Schreiben sowie visuokonstruktive Fertigkeiten werden überprüft. Im Bereich leichter kognitiver Beeinträchtigung (MCI) bzw. leichter demenzieller Syndrome sowie bei schwerer Demenz differenziert die MMSE nicht gut; für die Untersuchung im Hinblick auf eine leichte kognitive Beeinträchtigung (MCI) sind andere Verfahren besser geeignet. Im Bereich mittelschwerer demenzieller Syndrome ist die Korrelation mit anderen Assessmentinstrumenten gut, was zusammen mit der großen Verbreitung dafürspricht, die MMSE auch weiterhin zu verwenden. Die S3-Leitlinie Demenzen [AWMF Reg.Nr. 038-013] führt für die orientierende Schweregrad-Einteilung die Punktbereiche 20-26 (leichte Demenz), 10-19 (mittelschwer) und 0-9 (schwer) auf.

Die **MMSE-blind** (Tabelle 2.4.10) ist eine gekürzte Variante der Mini Mental State Examination (Items 1-22, maximal 22 Punkte) für Probanden mit Sehbehinderung mit vergleichbarer diagnostischer Güte wie die MMSE.

Die **Severe Mini Mental State Examination (SMMSE)** (Tabelle 2.4.11) ist ein praktikables, relativ schnell durchführbares Verfahren für Menschen mit einer schweren Demenzerkrankung. Das Maximum beträgt 30 Punkte, wer den MMSE-Test kennt und durchführt, wird sehr schnell mit dem SMMSE-Test vertraut sein.

Der **DemTect** (Tabelle 2.4.12) wurde in Deutschland entwickelt und evaluiert. Es liegt auch eine englischsprachige Version vor; eine verkürzte Fassung der englischsprachigen Version wird in Kanada in Zusammenhang mit der Fahreignung diskutiert. Folgende Aufgaben werden zur Überprüfung der kognitiven Funktionen eingesetzt: Wortliste mit unmittelbarer Abfrage (verbales Kurzzeit- bzw. Arbeitsgedächtnis), Zahlen umwandeln (kognitive Flexibilität, Lesen und Schreiben, Zahlenverarbeitung), „Supermarkt-Aufgabe“ mit Wortgenerierung (kognitive Flexibilität, Strategieanwendung, Wortfindung, kognitive Verarbeitungsgeschwindigkeit), Zahlenfolge rückwärts wiederholen (Arbeitsgedächtnis), verzögerte Abfrage der Wortliste (Einspeicherfähigkeit, Neugedächtnis). Die Rohwerte für die einzelnen Aufgaben müssen in Punkte umgerechnet werden. Dabei wird in zwei Altersklassen (unter bzw. ab 60 Jahre) unterschieden; darüber hinaus existiert keine Stratifizierung nach Alter, Ausbildung oder Geschlecht. Maximal 18 Punkte sind erreichbar; Ergebnisse ab 13 Punkten gelten als unauffällig [Kalbe 2004]. Im Bereich leichter kognitiver Einschränkungen ist der DemTect sensitiver als die MMSE.

Das **Montreal Cognitive Assessment (MoCA)** (Tabelle 2.4.13) wurde in Kanada als schnell durchzuführender Test entwickelt und evaluiert. Im Bereich leichter kognitiver Einschränkungen und leichter Demenzsyndrome ist das MoCA sensitiver als die MMSE. Es wurde weltweit bei verschiedenen neurokognitiven Erkrankungen untersucht, die Literatur ist sehr umfangreich. Es liegen Versionen in zahlreichen Sprachen vor, auch in Deutsch. Testformular und Manual auf Deutsch können auf der offiziellen MoCA-Internetseite kostenlos heruntergeladen werden (von Fachpersonal nach unkomplizierter kostenloser Registrierung).

Die folgenden kognitiven Funktionen (Domänen) werden geprüft: Aufmerksamkeit und Konzentration, Exekutivfunktionen, Gedächtnis, Sprache, visuokonstruktive Fähigkeiten, konzeptuelles Denken, Rechnen, Orientierung. Dabei sind 13 Aufgaben zu bewältigen: Trail-Making, Würfel zeichnen, Uhr-Zeichen-Test (mit max. 3 Punkten bewertet), Benennungsaufgabe, Wortliste (unmittelbare Reproduktion), Zahlenfolge rückwärts (Aufmerksamkeit, Arbeitsgedächtnis), Klopfen nur bei best. Buchstaben (Aufmerksamkeit, Inhibition), Subtraktionsaufgabe wie in der MMSE (Rechnen/Arbeitsgedächtnis), Satzwiederholung (Sprachverständnis), Wörter nennen, die mit F beginnen (phonematische Wortgenerierung), Unterschiede und Gemeinsamkeiten bei Wortpaaren (konzeptuelles Denken, Abstraktion), Wortliste (verzögerter Gedächtnisabruf), Orientierung Zeit/Ort. Maximal 30 Punkte sind zu erreichen. Personen mit maximal 12 Jahren Ausbildung erhalten einen zusätzlichen Punkt. Ursprünglich wurde ein Ergebnis von unter 26, inzwischen zur Verbesserung der diagnostischen Genauigkeit und zur Verringerung falsch positiver Ergebnisse alternativ ein Unterschreiten von 25 [Ciesielska 2016] oder 24 Punkten [Carson 2018] als Hinweis auf kognitive Störung betrachtet.

Es existieren Konversionstabellen zum Vergleich der Scores von DemTect, MMSE und MoCA (Tabellen 2.4.16, 2.4.17).

Mittlerweile sind - teilweise bereits mit deutscher Übersetzung - verschiedene Kurzformen sowie Versionen für Personen mit niedrigem Bildungsniveau, schwerer Seh- oder Hörbehinderung verfügbar.

Das **Montreal Cognitive Assessment – Hearing Impairment (MoCA-HI)** (Tabelle 2.4.14) wurde auf Deutsch übersetzt und validiert. Items, die ausreichende Hörfähigkeit voraussetzen, wurden durch adäquate visuelle Aufgaben ersetzt. Personen mit maximal 12 Jahren Ausbildung erhalten 2 zusätzliche Punkte. Der Test kann sowohl in Papierform als auch am Computer durchgeführt werden.

Das **Montreal Cognitive Assessment – Blind (MoCA-B)** [Wittich 2010] (deutsche Version auf der MoCA-Webseite verfügbar) wurde nicht validiert. Es ist identisch mit dem für die telefonische Erhebung bei Schlaganfall-Patient*innen validierten **T-MoCA** [Zietemann 2017]. Die eine ausreichende Sehfähigkeit voraussetzenden ersten vier Aufgaben des Original-MoCA wurden für den MoCA-B entfernt, wodurch sich die erreichbare Höchstpunktzahl auf 22 verringert. Die Entwicklung und Validierung einer Variante für Sehbeeinträchtigung (**MoCA-V**), bei der ähnlich dem Vorgehen beim MoCA-HI die ersten vier Aufgaben adäquat ersetzt werden, ist geplant.

Der **Test zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD)** (Tabelle 2.4.15) wurde in Deutschland entwickelt und evaluiert. Er ist im Bereich leichter kognitiver Einschränkungen und leichter Demenzsyndrome sensitiver als die MMSE. Die kognitiven Funktionen (Domänen) werden folgendermaßen geprüft: unmittelbare und verzögerte Reproduktion von sieben Wörtern, die die Patient*in zunächst laut vorliest (Gedächtnis), Fragen zur zeitlichen Orientierung (Datum, Jahreszeit) gepaart mit Kategorisierungsaufgaben (Zuordnung von Monaten zur aktuellen Jahreszeit), Befolgen von Anweisungen (Fähigkeit zu komplexeren, nicht automatisierten Handlungsabläufen), Uhr-Zeichen-Test nach Sunderland (konstruktive Praxie), kategoriale Wortflüssigkeits-Aufgabe (Tiere nennen) (semantisch-kategoriale Wortgenerierung). Die Depressivität wird auf einer 10er-Skala jeweils von Patient*in und Untersucher*in global eingeschätzt und die Summe aus beiden Werten gebildet. Im Demenzteil sind maximal 50 Punkte zu erreichen, ein Testergebnis von 35 oder weniger Punkten ist ein Hinweis auf eine demenzielle Erkrankung. Im Depressionsteil entsprechen mehr als 8 Punkte einer klinisch relevanten Symptomatik, 20 Punkte der stärksten Ausprägung depressiver Symptomatik.

Die **Global Deterioration Scale (GDS kog.)** (Tabelle 2.4.16) ist eine Fremdbeurteilungsskala, die eine Schweregradeinschätzung der Alzheimer-Demenz ermöglicht. Die Einstufung erfolgt meist auf der Basis eines klinischen Interviews über Alltagskompetenz und selbständige Versorgung der Person, was bei kognitiven Beeinträchtigungen die Anwesenheit von Angehörigen bzw. in die Pflege involvierten Personen voraussetzt. Die Skala erfordert eine gute Kenntnis von Demenzerkrankungen und kann für die Einschätzung der Schwere der Demenz zum Beispiel bei Pflegegeldgutachten, für Verlaufsbeobachtungen und die Aufklärung von Angehörigen über die zu erwartende Entwicklung sehr hilfreich sein.

Tab. 2.4.1: Kognition

Name des Instrumentes:	Six-Item Screener		
Kürzel:	SIS		
Mittlerer Zeitaufwand: Für Patient*in 1 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 1 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch: Ärzt*in, Pflegekraft, Ergotherapeut*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☐ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien: Nowroozpoor (2022): Zwei Scoping Reviews, eingeschlossen 45 Publikationen über die diagnostische Genauigkeit, 66 über die Praktikabilität des Assessments zur Detektion kognitiver Beeinträchtigung in der Notaufnahme. Von 27 Instrumenten war der SIS das mit dem geringsten Zeitbedarf und dem höchsten negativen prädiktiven Wert (Vergleichstest: MMSE) Krupp (2018) Studienpopulation n = 165, davon 57,6% kognitiv unauffällig, 80,9 ± 7,5 (56 – 97) J, 61,8% F, konsekutiv aufgenommen, stationäre Akutgeriatrie. Erfassung alltagsrelevanter Störung der Kognition: Bei Cut-off 5/4 Sensitivität 100%, Spezifität 71%, AUC 0,937, positiver prädiktiver Wert 0,39. Bei Cut-off 4/3 positiver prädiktiver Wert 0,54. Retest-Reliabilität 0,705, interne Konsistenz Cronbachs α 0,821, Korrelation mit abschließendem Urteil des verblindeten Geriatriers -0,729, mit MMSE 0,677, mit Uhrentest nach Shulman -0,478. Änderungssensitivität: partielles η^2 0,25 (Werte aus der Validierungsstudie der deutschen Übersetzung). Carpenter (2011): Studienpopulation n = 319 über 65-jährige Patienten, mittleres Alter 76 J, 58% F, 58% afroamerikanisch, vital stabile Patient*innen jedweder Diagnose ohne sedierende Medikation in der Notaufnahme. Der SIS erkannte bei Cut-off 5/4 74% der Patient*innen mit MMSE <24 (AD8 63%, wenn Angehörige befragt werden konnten, 37%, wenn Patient*innen interviewt wurden). Winter (2005): Sensitivität 94%, Spezifität 86%, negativer prädiktiver Wert 0,98, AUC 0,96			
Begründung der Aufnahme in die Leitlinie: Mit minimalem Zeitaufwand materialfrei durchführbar, Visus ohne Einfluss, hohe Akzeptanz.			
Limitationen / Forschungsbedarf: Forschungsbedarf besteht für die Änderungssensitivität wegen niedriger Fallzahl an Proband*innen, die sich im Beobachtungszeitraum in der MMSE und dem Uhrentest nach Shulman gleichsinnig veränderten.			

Literatur:

Callahan CM, Unverzagt FW, Hui SL et al. (2002) Six-item screener to identify cognitive impairment among potential subjects for clinical research. Med Care 40:771–781

Carpenter CR, DesPain B, Kealing TN et al. (2011) The Six-Item Screener and AD8 for the Detection of Cognitive Impairment in Geriatric Emergency Department Patients. Ann Emerg Med 57(6): 653–661

Deutsche Version in:

Krupp S, Seebens A, Kasper J. et al. (2018) Validierung der deutschen Fassung des Six-Item Screeners. Kognitiver Kurztest mit breiten Anwendungsmöglichkeiten. Z Gerontol Geriat 51: 275–281. DOI: 10.1007/s00391-016-1177-z

Nowroozpoor A, Dussetschleger J, Perry W et al. (2022) Detecting Cognitive Impairment and Dementia in the Emergency Department: A Scoping Review. J Am Med Dir Assoc 23(8): 1314.e31–1314.e88. DOI:10.1016/j.jamda.2022.03.019

Sterling MR, Ringel JB, Safford MM et al. (2024) Trajectory of cognitive decline after incident heart failure hospitalization: Findings from the REGARDS study. J Am Heart Assoc 13(17):e032986. DOI: 10.1161/JAHA.123.032986

Wilber ST, Lofgren SD, Mager TG et al. (2005) An evaluation of two screening tools for cognitive impairment in older emergency department patients. Acad Emerg Med 12:612–616

Xue J, Chiu HF, Liang J et al. (2018) Validation of the Six-Item Screener to screen for cognitive impairment in primary care settings in China, Aging & Mental Health, 22:453-457. DOI: 10.1080/13607863.2017.1280768

Tab. 2.4.2: Kognition

Name des Instrumentes:	Uhr-Zeichen-Test nach Shulman		
Kürzel:	–		
Mittlerer Zeitaufwand: Für Patient*in 3 min. / für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min.			
Lernaufwand für Untersucher*in:			
☒ leicht		☐ mittel	☐ anspruchsvoll
Durchführung durch: Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung		☒ mittel	☐ schwer
Gütekriterien: Sensitivität und Spezifität jeweils 85% [Shulman 2000]. Vergleichende Analyse der verschiedenen Auswertesysteme des CDT einer deutschen Memory-Clinic ergab durchwegs hohe Interrater-Reliabilität von 0,82 bis 0,94 (Shulman-Methode: 0,89 – Watson-Methode 0,90) [Watson 1993], die höchste diagnostische Sicherheit erreichte die Methode nach Shulman (Sensitivität 81% und Spezifität 79%) (Auswertung nach Watson: Sensitivität			

56%, Spezifität 80%) [Schramm 2002]. Auch in Kombination mit MMSE erbrachte die Auswertung des CDT nach Shulman die höchste Sensitivität von 90% (Watson-Auswertung 84%) [Watson 1993]. Ähnliche Ergebnisse in vergleichender Analyse einer australischen Memory-Clinic: bei der Auswertung nach Shulman die höchste Genauigkeit gegenüber 4 weiteren (ROC 0,79, im Vergleich Watson-Methode 0,65) [Storey 2001]. Eine weitere australische Studie mit nur 28 Patienten favorisiert ebenfalls die Auswertung nach Shulman [Brodaty 1997].

Begründung der Aufnahme in die Leitlinie:

Höhere diagnostische Sicherheit als Uhrentest nach Watson.

Limitationen / Forschungsbedarf:

Bei Menschen mit schweren Beeinträchtigungen des Sehens oder der Feinmotorik der Schreibhand ist der Uhr-Zeichen-Test nicht durchführbar.

Literatur:

Bertol Graeff D, Maldaner Lui J, Dal Prá Zucco N, Sant'Anna Alves AL, Forcelini CM, Dalmolin BM (2021) Clock drawing test: comparison between the Pfizer and the Shulman systems. *Dement Neuropsychol* 15(4): 480-484

Brodaty H, Moore CM (1997) The Clock Drawing Test for dementia of the Alzheimer's type: A comparison of three scoring methods in a memory disorders clinic. *Int J Geriatr Psychiatry* 12:619-627

Claus CC, Staekenborg SS, Verweij HW et al. (2023) The clock drawing test is an important contribution to the Mini Mental State Examination in screening for cognitive impairment. *Int J Geriatr Psychiatry* 38(4)e5914. DOI: 10.1002/gps.5914

Mainland BJ, Amodeo S, Shulman KI (2014) Multiple clock drawing scoring systems: simpler is better. *Int J Geriatr Psychiatry* 29:127-136

Mendez MF, Ala T, Underwood KL (1992) Development of scoring criteria for the clock drawing task in Alzheimer's disease. *J Am Geriatr Soc* 40:1095-1099

Park J, Jeong E, Seomun G (2018) The clock drawing test: A systematic review and metaanalysis of diagnostic accuracy. *J Adv Nurs* 74(12):2742-2754

Ribera AAV, Lino TB, Rodrigues NO, Ansai JH, de Andrade LP, Christofolletti G (2025) Has the Clock Drawing Test been left aside with the replacement of analog clocks by smartphones? *Dement Neuropsychol* 19 <https://doi.org/10.1590/1980-5764-DN-2024-0178>

Schramm U, Berger G, Müller R et al. (2002) Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population., *Int J Geriatr Psychiatry* 17:254-260

Shulman KI, Shedletsky R, Silver IL (1986) The challenge of time: clock-drawing and cognitive function in the elderly. *Int J Geriatr Psychiatry* 1:135-140

Shulman KI, Gold DP, Cohen CA, Zuccherro CA (1993) Clock-drawing and dementia in the community: A longitudinal study. *Int J Geriatr Psychiatry* 8:487-496

Shulman KI (2000) Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry* 15:548-561

Spenciare B, Alves H, Charchat-Fichman H (2017) Scoring systems for the Clock Drawing Test. A historical review. *Dement Neuropsychol* 11:6-14

Storey JE, Rowland JT, Basic D, Conforti DA (2001) A comparison of five clock scoring methods using ROC (receiver operating characteristic) curve analysis. *Int J Geriatr Psychiatry* 16:394-399

Tan LP, Herrmann N, Mainland BJ, Shulman K (2015) Can clock drawing differentiate Alzheimer's disease from other dementias?, *Int Psychogeriatr* 27:1649-1660

Tab. 2.4.3: Kognition

Name des Instrumentes:	Uhr-Zeichen-Test nach Sunderland		
Kürzel:	–		
Mittlerer Zeitaufwand: Für Patient*in 3 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min			
Lernaufwand für Untersucher*in:			
O leicht	x mittel	O anspruchsvoll	
Durchführung durch: Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
x leicht/Früherkennung	x mittel	O schwer	
Gütekriterien: Schramm (2002): Bei Cut-off 5/6 Sensitivität 56% und Spezifität 91%, bei Cut-off 8/9 Sensitivität 85% und Spezifität 80 %. Brodaty (1997): Bei Cut-off von 5/6 Sensitivität 78% und Spezifität 96%. Sunderland (1989): Studienpopulation n = 67 mit M. Alzheimer, 67,7 ± 9,6 J und n = 83 in der kognitiv gesunden Kontrollgruppe, 73,6 ± 7,9 J, ambulantes Setting. Interrater-Reliabilität 0,86 (p < 0,001). Alzheimer-Patient*innen MW 4,9 ± 2,7 Punkte, Klient*innen der kognitiv gesunden Kontrollgruppe MW 8,7 ± 1,1 Punkte (p < 0,001). Korrelationen zwischen den Uhrzeichnungen und der Global Deterioration Scale 0,56), der Dementia Rating Scale 0,59 und der Blessed Dementia Rating Scale 0,51 (p < 0,001) durch drei verschiedene Untersucher*innen unter Einbezug des klinischen Eindrucks. Cut-off 5/6 empfohlen.			
Begründung der Aufnahme in die Leitlinie: Enthalten in Mini-Cog, „Schnellem Uhren-Dreier“ und TFDD, keine Unterlegenheit gegenüber Uhrentest nach Watson, evtl. höhere Änderungssensitivität.			
Limitationen / Forschungsbedarf: Bei Menschen mit schweren Beeinträchtigungen des Visus oder der Feinmotorik der Schreibhand ist der Uhr-Zeichen-Test nicht durchführbar.			
Literatur: Brodaty H, Moore CM (1997) The Clock Drawing Test for dementia of the Alzheimer's type: A comparison of three scoring methods in a memory disorders clinic. Int J Geriatr Psychiatry 12:619-627 Maestri G, Nicotra A, Pomati S et al. (2023) Cultural influence on clock drawing test: A systematic review. J Int Neuropsychol Soc 29:704–714. DOI:10.1017/S1355617722000662 Park JK, Jeong EH, Seomun GA (2018) The clock drawing test: A systematic review and meta-analysis of diagnostic accuracy. J Adv Nurs, 74(12): 2742-2754 Schneider F (Hrsg.) (2008) Klinikmanual Psychiatrie, Psychosomatik und Psychotherapie. Springer Verlag Berlin Heidelberg			

Schramm U, Berger G, Müller R et al. (2002) Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population., Int J Geriatr Psychiatry, 17:254-260

Sunderland T, Hill J, Mellow A et al. (1989) Clock drawing in Alzheimer’s disease: A novel measure of dementia severity. J Am Geriatr Soc 37:725-729

Tab. 2.4.4: Kognition

Name des Instrumentes:	Uhren Ergänzen (Clock Completion) nach Watson		
Kürzel:	CC		
Mittlerer Zeitaufwand:			
Für Patient*in 3 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien:			
Tuokko (2000): Studienpopulation n = 493, 79 ± 7 J, 53% F. Vergleich von 5 Uhr-Zeichen-Tests, bei dem der nach Watson die niedrigste diagnostische Sicherheit erreichte (AUC Watson 0,67, Shulman 0,79).			
Watson (1993): Studienpopulation n = 104, 76 ± 8 J, allgemeinmedizinische Praxis und Krankenhaus, Sens. 87%, Spez. 82%, Interrater-Reliabilität 0,85-0,95, Retest-Reliabilität 0,25 – 0,43 (Korrelationen nach Spearman), Korrelation mit dem MMSE für Zeichenaufgabe 0,35 (Kappa).			
Begründung der Aufnahme in die Leitlinie:			
Von der AGAST empfohlenes Instrument.			
Limitationen / Forschungsbedarf:			
Bei Menschen mit schweren Beeinträchtigungen des Sehens oder der Feinmotorik der Schreibhand ist der Test nicht valide durchführbar.			
Literatur:			
Die Bewertungsvorlage ist unter https://www.kcgeriatrie.de/assessments-in-der-geriatrie/assessmentbereiche/kognition verfügbar (Zugriff: 17.11.2025)			
Agrell, B, Dehlin, O (1998). The clock-drawing test. Age Ageing 27: 399-403			
Arbeitsgruppe Geriatrisches Assessment (AGAST) (1997) Geriatrisches Basisassessment. Handlungsanleitungen für die Praxis. MMV Medizin Verlag München, 2. Auflage			
Ferruci, L et al. (1996). Does the Clock Drawing Test predict cognitive decline in older persons independent of the MMSE. JAGS 44: 1326-1331			

Schweer R, Naumann C (2001) Der Uhrentest. Eine methodische Überprüfung hinsichtlich seines Einsatzes im geriatrischen Basisassessment. Zeitschrift für Gerontopsychologie & -psychiatrie 14:123–136

Specht-Leible N (1999) Kognitive Defizite einfach von der Uhr ablesen. Geriatrie Praxis 12:37–38

Tuokko H, Hadjistavropoulos T, Rae S, O'Rourke N (2000) A Comparison of Alternative Approaches to the Scoring of Clock Drawing. Arch Clin Neuropsychology 15(2):137-148

Van Hout, H, Berhout, S (1999). Inter-rater Reliability of he clock-drawing test. Age Ageing 28: 327-328

Watson H, Arfken CL, Birge SJ (1993) Clock completion: An objective Screening test for dementia. J Am Geriatr Soc 41:1235–1240

Tab. 2.4.5: Kognition

Name des Instrumentes:	Mini-Cog		
Kürzel:	–		
Mittlerer Zeitaufwand: Für Patient*in 3 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 5 min			
Lernaufwand für Untersucher*in:			
☒ leicht		☐ mittel	☐ anspruchsvoll
Durchführung durch: Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung		☒ mittel	☐ schwer
Gütekriterien: Stewart (2022): Zusammenfassung von relevanten Cochrane-Reviews. Die Autor*innen schlussfolgern, in der Primärversorgung erreiche der Mini-Cog keine ausreichende prädiktive Validität im Hinblick auf eine Demenzentwicklung. Chan (2021) (Cochrane Database of Systematic Reviews): 3 Studien mit historischen Daten aus der Sekundärversorgung (ambulante Krankenhauspatienten, Memory Clinic) Sensitivität 60% - 87%, Spezifität 65% - 100%; nur 55,3% der Teilnehmenden wurden in die Datenanalyse einbezogen, hohes Risiko von Auswahl-Bias. Fage (2021) (Cochrane Database of Systematic Reviews): Im Kontext der gemeindenahen Versorgung in drei randomisiert-kontrollierten Studien Sensitivität 0,76 – 0,99, Spezifität 0,83 – 0,93; Studienqualität limitiert, Studien zur Genauigkeit im Vergleich zu anderen kognitiven Assessmentverfahren erforderlich. Im Kontext der Primärversorgung Sensitivität 76% – 100%, Spezifität 27% – 85%; Studienqualität limitiert; unzureichende Evidenz für Empfehlung zum Assessment im Vergleich zur jeweils eingesetzten vertieften Evaluation durch Demenz-Expert*innen. Borson (2003): Sekundäre Auswertung der Daten einer in den frühen 1990er Jahren in den USA durchgeführten Studie. Eingeschlossen wurden >65-jährige Teilnehmende (n = 1119),			

mit Prävalenz von Demenz gemäß DSM von 6,4% (n = 74), fraglich kognitiv eingeschränkte Personen wurden von der Analyse ausgeschlossen. Die Komponenten des Mini-Cog (wobei der Uhrentest in der Original-Studie anders bewertet worden war) wurden aus den Daten zusammengestellt. 73,1 ± 6,0 J, 54,6% F, amb., Median formaler Bildung 12 J. Sensitivität 76% (MMSE mit Cut-off 25/24 79%, 30-minütige neuropsycholog. Batterie 75%), Spezifität 89%, (MMSE 88%, neuropsycholog. Batterie 90%).

Begründung der Aufnahme in die Leitlinie:

Erweitert die mit dem Uhrentest erfassten kognitiven Domänen um das Kurzzeitgedächtnis. Bei der Bewertung ist zu beachten, dass auch eine isolierte Beeinträchtigung der Uhren-Aufgabe zur weiteren Abklärung führen sollte.

Limitationen / Forschungsbedarf:

Bei Menschen mit schweren Beeinträchtigungen des Sehens oder der Feinmotorik der Schreibhand ist der Test nicht durchführbar.

Literatur:

Abayomi SN, Sritharan P, Yan E et al. (2024) The diagnostic accuracy of the Mini-Cog screening tool for the detection of cognitive impairment - A systematic review and meta-analysis. PLoS ONE 19(3): e0298686

Borson S, Scanlan JM, Chen P, Ganguli M (2003) The Mini-Cog as a screen for dementia: Validation in a population-based sample. JAGS 51:1451–1454

Chan CCH, Fage BA, Burton JK et al. (2021) Mini-Cog for the detection of dementia within a secondary care setting. Cochrane Database of Systematic Reviews, 7.

DOI: 10.1002/14651858.CD011414.pub3.

Fage BA, Chan CCH, Gill SS et al. (2021) Mini-Cog for the detection of dementia within a community setting. Cochrane Database of Systematic Reviews, 7, DOI: 10.1002/14651858.CD010860.pub3

Gomes-Osman J, Borson S, Toro-Serey C, Banks R, Ciesla M, Jannati A, Morrow WI, Swenson R, Libon D, Bates D, Showalter J, Tobyne S, Pascual-Leone A (2024) Digital Clock and Recall: a digital, process-driven evolution of the Mini-Cog. Front Hum Neurosci. 18:1337851.

DOI: 10.3389/fnhum.2024.1337851

Ismail Z, Rajji TK, Shulman KI (2010) Brief cognitive screening instruments: an Update. Int J Geriatr Psychiatry 25:111-120

Seitz DP, Chan CCH, Newton HT et al. (2021) Mini-Cog for the detection of dementia within a primary care setting. Cochrane Database of Systematic Reviews, 7

DOI: 10.1002/14651858.CD011415.pub3

Stewart C, Stewart J, Stubbs S et al. (2022) Mini-Cog, IQCODE, MoCA, and MMSE for the Prediction of Dementia in Primary Care. Am Fam Physician 105: 590-592

Tiway N, Treggiari MM, Yanez ND et al. (2021) Agreement Between the Mini-Cog in the Preoperative Clinic and on the Day of Surgery and Association With Postanesthesia Care Unit Delirium: A Cohort Study of Cognitive Screening in Older Adults. Anesthesia & Analgesia: 132:1112-1119 DOI: 10.1213/ANE.0000000000005197

Yang Q, Yang H, Long W, et al. (2025) Rapid cognitive assessment: Accuracy and discriminant validity of Mini-Cog and process-based Clock Drawing Test. J Alzheimers Dis 104: 200-208

Tab. 2.4.6: Kognition

Name des Instrumentes: Kürzel:	„Schneller Uhren-Dreier“ SUD	
Mittlerer Zeitaufwand: Für Patient*in 3 min. / für Untersucher*in (inkl. Vor- u. Nachbereitung) 5 min.		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch: Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien: Kamenski (2009): prospektive Validierungsstudie mit 107 Patient*innen, 86 mit vollständigem Studienprotokoll, MW 77,3 ± 6,9 J; 63% F; ambulantes Setting →Sensitivität 85% (95% CI: 71%, 98%), Spezifität 58% (95% CI: 46%, 71%). PPW 0,47 (95% CI: 0,33, 0,61), NPW 0,90 (95% CI: 0,80, 0,99).		
Begründung der Aufnahme in die Leitlinie: Vom Mini-Cog abgeleitetes, für die Primärversorgung vereinfachtes Instrument.		
Limitationen / Forschungsbedarf: Bei Menschen mit schweren Beeinträchtigungen des Sehens oder der Feinmotorik der Schreibhand ist der Test nicht durchführbar.		
Literatur: Geschke K, Weyer-Elberich V, Binder H, Fellgiebel A (2020) Screening for Cognitive Impairment in Hospitalized Older Patients. GeroPsych https://doi.org/10.1024/1662-9647/a000253 Ivemeyer D, Zerfaß R (2006) Demenztests in der Praxis. Ein Wegweiser. Verlag Urban & Fischer, 2. Auflage München – Jena S.85 Kamenski G, Dorner T, Lawrence K et al (2009) Detection of dementia in primary care: comparison of the original and a modified Mini-Cog Assessment with the Mini-Mental State Examination. Ment Health Fam Med 6:209-217 Krupp S (2017) Geriatrisches Assessment. In Willkomm M (Hrsg.) Praktische Geriatrie. 2. Auflage. Georg Thieme Verlag Stuttgart S.51 <u>Strotzka S, Psota G, Sepandj A (2003) Uhrentest in der Demenzdiagnostik – Auf der Suche nach der verlorenen Zeit. Psychopraxis 04/03:16-24</u>		

Tab. 2.4.7: Kognition

Name des Instrumentes:	Brief Alzheimer Screen		
Kürzel:	BAS		
Mittlerer Zeitaufwand:			
Für Patient*in 4 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 6 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☐ mittel	☐ schwer	
Gütekriterien:			
Carpenter (2011): In der Notaufnahme (Emergency Department) Sensitivität 95% in der Detektion einer kognitiven Dysfunktion.			
<u>Mendiondo (2003)</u> : Daten aus einer Studie zu CERAD wurden genutzt, um Items für den BAS auszuwählen und die Gütekriterien mathematisch zu bestimmen (ohne prospektiven Einsatz des so konstruierten Instruments). Studienpopulation: ambulante Probanden nach Ausschluss diverser Erkrankungen als möglicher sekundärer Demenzursachen, n = 748, davon n=342 mit leichtgradiger Alzheimer-Demenz bei einem MMSE von mindestens 20 Punkten, 72,8 ± 7,7 J, 66 % F, n=406 mit altersgerechter Kognition, 59% F, 68,6 ± 7,9 J. Bei Cut-off 27/26 Sensitivität von 99% und Spezifität von 87%, bei Cut-off 23/22 Sensitivität 90% und Spezifität 99%. 23-26 Punkte passen zu einer leichten kognitiven Störung, der Befund ist kontrollbedürftig.			
Begründung der Aufnahme in die Leitlinie:			
Auch bei Visus-, Hör- oder feinmotorischer Störung einsetzbarer kognitiver Test, im Vergleich zur MMSE weniger Zeitbedarf, erkennt frühe Alzheimer-Krankheit sensitiver.			
Limitationen / Forschungsbedarf:			
Der Test eignet sich nicht zur Schweregradeinschätzung einer Demenz. Bislang fehlt eine validierte deutsche Übersetzungsversion, zumal für das Rückwärtsbuchstabieren des in der amerikanischen Version aus fünf Buchstaben bestehenden Wortes WORLD eine deutsche Entsprechung mit fünf Buchstaben fehlt bzw. noch nicht konsentiert ist.			
Literatur:			
Carpenter CR, Bassett ER, Fischer GM et al. (2011) Four Sensitive Screening Tools to Detect Cognitive Dysfunction in Geriatric Emergency Department Patients: Brief Alzheimer's Screen, Short Blessed Test, Ottawa 3DY, and the Caregiver-completed AD8. Acad Emerg Med 18:374-384 DOI:10.1111/j.1553-2712.2011.01040.x			
<u>Mendiondo MS, Ashford JW, Kryscio RJ, Schmitt FA (2003) Designing a Brief Alzheimer Screen (BAS). J Alzheimers Dis 5:391-398</u>			
https://alzfdn.org/memory-screening/what-is-memory-screening/			

Tab. 2.4.8: Kognition

Name des Instrumentes:	Bamberger Demenz Screening Test		
Kürzel:	BDST		

Mittlerer Zeitaufwand:
Für Patient*in 8 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 12min

Lernaufwand für Untersucher*in:

☒ leicht	☐ mittel	☐ anspruchsvoll
---	------------------------------	-------------------------------------

Durchführung durch:
Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal.

Geeignet für Schweregrad:

☒ leicht/Früherkennung	☒ mittel	☐ schwer
---	---	------------------------------

Gütekriterien:
Trapp (2021): Studienpopulation n = 1295, 76 ± 8 J, 61% F, Patient*innen, die zur neuropsychologischen Testung einschließlich CERAD-plus aus geriatrischen Krankenhaus-Abteilungen vorgestellt wurden. Interne Konsistenz: Cronbachs α 0,79. Als Vergleichsstandard diente das Urteil des bezüglich des für den BDST verblindeten Teams, dabei wiesen 136 Patient*innen keine Hinweise auf kognitive Störung auf. Bei Cut-off von 38/37 für den 16 Items umfassenden BDST (max. 50 Punkte) Sensitivität in der Detektion leichter Alzheimer-Demenz 94% (MMSE hier 79%), Spezifität 82% (MMSE 88%), für andere Formen leichter Demenz (n = 364) Sensitivität 91%, Spezifität 82%. Für die Detektion moderater Demenz: BDST bei Cut-off von 32/31 Sensitivität und Spezifität 100%, BDST short form (Items 1 – 9, max. 18 Punkte) bei Cut-off 16/15 Sensitivität 94%, Spezifität 93%, BDST ultra-short form (Items 1 – 5, max. 10 Punkte) bei Cut-off 10/9 Sensitivität 85%, Spezifität 86%. Korrelation BDST/CERAD-plus r = 0,77, BDST Short Form/CERAD-plus r = 0,56, BDST Ultra Short/DERAD-plus r = 0,42.

Begründung der Aufnahme in die Leitlinie:
Ohne Material durchführbar, in der Früherkennung sensitiver als MMSE, je nach Intention und in (Ultra-) Kurzversion erzieltm Wert kann der Umfang des Tests reduziert werden.

Limitationen / Forschungsbedarf:
Die o.g. Daten zur Validität entstammen einer in Bezug auf die Prävalenz von Demenz ausgelesenen Stichprobe. Die Durchführung des Tests muss ohne Störgeräusche erfolgen.

Literatur:
Trapp W, Weisenberger B, Düclos D. et al. (2015) The Bamberg Dementia Screening Test (BDST) – first evidence regarding the diagnostic usability of a “true bedside” test for geriatric inpatients. Z Neuropsychol 26:161-170 BMC Med. 2021; 19: 65.
Trapp W, Röder S, Heid A et al. (2021) Sensitivity and specificity of the Bamberg Dementia Screening Test's (BDST) full and short versions: brief screening instruments for geriatric patients that are suitable for infectious environments BMC 19:65 DOI: 10.1186/s12916-021-01927-4

Tab. 2.4.9: Kognition

Name des Instrumentes:	Mini Mental State Examination		
Kürzel:	MMSE		
Mittlerer Zeitaufwand: Für Patient*in 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 14 min			
Lernaufwand für Untersucher*in:			
☐ leicht	☒ mittel	☐ anspruchsvoll	
Durchführung durch: Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☐ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien: Pinto (2019): In einem systematischen Review kommen die Autor*innen zu dem Schluss, dass MMSE und MoCA zur Demenzdetektion gleich geeignet sind, MoCA der MMSE zur Diagnose eines MCI aber überlegen ist. Arevalo-Rodriguez (2015), Cochrane Review: Sensitivität zwischen 27% [Devanand 2008] und 89% [Buchhave 2008], Spezifität zwischen 32% [Buchhave 2008] und 90% [Devanand 2008], in der Summe gemäß ROC Analyse Sensitivität 54% bei einer Spezifität von im Median 80%. Folstein (1975): 206 stationäre Patient*innen einer psychiatrischen Klinik, 63 Senior*innen ohne Verdacht auf Demenz als Vergleichsgruppe. Retest-Reliabilität 0,887 (Pearson Korrelation). Begründung der Aufnahme in die Leitlinie: Von der AGAST empfohlenes, international breit eingesetztes Instrument, zur Abbildung des Verlaufs einer demenziellen Entwicklung einsetzbar. Limitationen / Forschungsbedarf: Geringe Sensitivität im Bereich der Früherkennung von dementiellen Entwicklungen, besonders bei Personen mit hohem prämorbidem Intelligenzniveau. Die MMSE vermag sowohl im oberen als auch unteren Wertebereich nicht gut zu differenzieren. Literatur: Arevalo-Rodriguez I, Smailagic N, Roqué et al. (2015) Mini-Mental State Examination (MMSE) for the detection of Alzheimer’s disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database of Systematic Reviews Issue 3. Art. No.: CD010783. DOI: 10.1002/14651858.CD010783.pub2. Bernini S, Ballante E, Picascia M et al. (2025) Equating conversion norms for the Mini-Mental State Examination and Montreal Cognitive Assessment in healthy subjects and patients with neurodegenerative disorders. Exp Gerontol 205:112756 DOI: 10.1016/j.exger.2025.112756 Brogaard NJ, Hahn-Pedersen JH, Gundgaard J et al. (2024) Mapping between the Mini-Mental State Examination and the Clinical Dementia Rating scale for Patients with Alzheimer’s Disease: An			

External Validation Study. Alzheimer's Dement 20 (Suppl. 7) <https://doi.org/10.1002/alz.087285>

Creavin ST, Wisniewski S, Noel-Storr AH et al. (2016) Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. Cochrane Database of Systematic Reviews 2016, Issue 1. Art. No.: CD011145. DOI: 10.1002/14651858.CD011145.pub2.

Folstein MF, Folstein SE, Mc Hugh PR (1975) „Mini-Mental State“ – A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189-198

Pinto T, Machado L, Bulgacov T et al. (2019) Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's Disease (AD) in the elderly? International Psychogeriatrics 31: 491-504

Salis F, Costaggu D, & Mandas A (2023) Mini Mental State Examination: Optimal cut-off levels for mild and severe cognitive Impairment. Geriatrics 8:12. DOI: 10.3390/geriatrics8010012

Strotzka, S (2004) 30 Jahre Mini-Mental State Examination – Eine kritische Würdigung. Psychopraxis 5:24-34

Trzepacz PT, Hochstetler H, Wang S et al. (2015) Alzheimer's disease neuroimaging initiative relationship between the Montreal Cognitive Assessment and Mini-mental State Examination for assessment of mild cognitive impairment in older adults. BMC Geriatr 15:107

Wang G, Estrella A, Hakim O, Milazzo P et al. (2022) Mini-Mental State Examination and Montreal Cognitive Assessment as Tools for Following Cognitive Changes in Alzheimer's Disease Neuroimaging Initiative Participants. J Alzheimers Dis 90(1):263-270

Tab. 2.4.10: Kognition

Name des Instrumentes:	Mini Mental State Examination blind	
Kürzel:	MMSE-blind	
Mittlerer Zeitaufwand:		
Für Patient*in 8 min / für Untersucher (inkl. Vor- u. Nachbereitung) ca. 10 min		
Lernaufwand für Untersucher*in:		
☐ leicht	☒ mittel	☐ anspruchsvoll
Durchführung durch:		
Ärzt*in, Psycholog*in		
Geeignet für Schweregrad:		
☐ leicht/Früherkennung	☒ mittel	☒ schwer
Gütekriterien:		
Riedel-Heller (2000): Leipzig Longitudinal Study of the Aged (LEILA 75+): n = 88, MW 86,4 ± 5,9 J; 86,4% F, ambulantes Setting. Cut-off variiert nach Alter und Bildung zwischen 16/15 (nach Busse bei Personen ab 80 Jahren mit niedriger Bildung) und 20/19 Punkten (nach Reischies bei Personen unter 85 Jahren mit hoher Bildung) bei hoher Sensitivität (91%-100%) und Spezifität (80%-100%) in Bezug auf Demenz unterschiedlicher Genese. Retest-Reliabilität: Wilcoxon-Test, z = -2,43, p =0,015		

Begründung der Aufnahme in die Leitlinie:
Alternative zur kompletten MMSE für Personen mit schwerem Visusverlust.

Limitationen / Forschungsbedarf:
Beim MMSE-blind besteht noch Forschungsbedarf hinsichtlich der Testgütekriterien.

Literatur:
Busse A, Sonntag A, Bischkopf J et al. (2002) Adaptation of dementia screening for vision-impaired older persons. Administration of the Mini-Mental State Examination (MMSE). J Clin Epidemiol 55:909-915
Reischies FM, Geiselman B (1997) Age-related cognitive decline and vision impairment affecting the detection of dementia syndrome in old age. Br J Psychiatry 171:449-451
Riedel-Heller SH, Schork A, Matschinger H, Angermeyer MC (2000) Recruitment procedures and their impact on the prevalence of dementia: results from the Leipzig Longitudinal Study of the Aged (LEILA 75+). Neuroepidemiology 19:130-140

Tab. 2.4.11: Kognition

Name des Instrumentes:	Severe Mini Mental State Examination		
Kürzel:	SMMSE		
Mittlerer Zeitaufwand:			
Für Patient*in 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 14 min			
Lernaufwand für Untersucher*in:			
☐ leicht	☒ mittel	☐ anspruchsvoll	
Durchführung durch:			
Ärzt*in, Psycholog*in			
Geeignet für Schweregrad:			
☐ leicht/Früherkennung	☐ mittel	☒ schwer	
Gütekriterien:			
Strotzka (2005): Validierung einer deutschen Version, 52 Patient*innen, MW 81 J; 65% F, ambulant Setting. Korrelation zwischen MMSE und SMMSE 0,87 (MMSE 20-12 Punkte / SMMSE 28,2 ± 2,5 Punkte; MMSE 11-0 Punkte / SMMSE 18,3 ± 7,9 Punkte).			
Harrell (2000): 182 Patient*innen, MW 73,3 ± 7,7 J; 65% F; stationäres Setting (Alzheimer-Zentrum und Neurologie). Interrater-Reliabilität 0,99 (p < 0,0001).			
Der Test wurde auch u. a. in Italienisch [Bianchetti 2002], Portugiesisch [Wajman 2014], Baskisch [Buiza 2018] und Griechisch [Mougias 2018] validiert.			
Begründung der Aufnahme in die Leitlinie:			
Bei gegebener Indikation Alternative für schwer kognitiv beeinträchtigte Patient*innen, insbesondere bei Überforderung mit Abbruch anspruchsvollerer Tests.			

Limitationen / Forschungsbedarf:

Der SMMSE-Test wurde speziell für Menschen mit schwerer Demenzerkrankung entwickelt, für den mittelschweren Demenzbereich sind die Aufgaben zu leicht.

Literatur:

Deutsche Version auf der Homepage der Deutschen Gesellschaft für Gerontopsychiatrie und –psychotherapie e.V. (DGGPP, www.dggpp.de) sowie auf der Homepage der Österreichischen Gesellschaft für Alterspsychiatrie und –psychotherapie (ÖGAPP, www.alterspsychiatrie.at)

Bianchetti A, Scala V, Cipriani G, Trabucchi M (2003) The Severe Mini Mental State Examination (SMMSE): Validation of the Italian version. *Res Pract Alzh Dis: Severe Dementia* 8:39-45

Buiza C, Yanguas J, Zulaica A et al. (2018) Evaluación cognitiva global en personas vascoparlantes con demencia en estadios avanzados. Validación al euskera de la escala *Severe Mini Mental State Examination* (SMMSE-eus). *Rev Esp Geriatr Gerontol* 53:208-212

Harrell LE, Marson D, Chatterjee A et al (2000) The Severe Mini-Mental State Examination: a new neuropsychologic instrument for bedside assessment of severely impaired patients with Alzheimer disease. *Alz Dis Assoc Dis* 14:168-175 DOI: 10.1097/00002093-200007000-00008

Mougias A, Christidi F, Kiosterakis G, Messinis L, Politis A (2018) Dealing with severe dementia in clinical practice: A validity and reliability study of Severe Mini-Mental State Examination in Greek Population. *Int J Geriatr Psychiatry* 2018 Jun 1. doi: 10.1002/gps.4915Strotzka S, Sepandj A, Psota G (2005) Severe Mini Mental State Examination – Neuropsychologie für schwer demente Menschen. *Psychopraxis* 6:10-15

Wajman JR, de Oliveira FF, Rizek Schultz R, de Medeiros Correia Marin S, Bertolucci PH (2014) Educational bias in the assessment of severe dementia: Brazilian cutoffs for severe Mini-Mental State Examination. *Arq Neuropsiquiatr* 72(4):273-277

Tab. 2.4.12: Kognition

Name des Instrumentes:	DemTect	
Kürzel:	—	
Mittlerer Zeitaufwand:		
Für Patient*in 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 12 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien:		
Ozer (2016): Im Rahmen eines systematischen Reviews zur Detektion eines amnestischen MCI geben die Autor*innen die Sensitivität mit 85% und die Spezifität des DemTect mit 86%		

an.

Kohn (2007): n = 83, 57,6 ± 14,7 J, stationäre Patient*innen Neurologie mit V. a. kognitive Störung, davon 41 kognitiv beeinträchtigt gemäß Expertenurteil, 42 unbeeinträchtigt gemäß zweier neuropsychologischer Tests. Bei Cut-off 13/12 Punkten Sensitivität 97,6%, Spezifität 92,9% für wahrscheinliche kognitive Beeinträchtigung (bei Cut-off 25/24 Punkten in der MMSE Sensitivität 46,3%, Spezifität 100%).

Kalbe (2005): Für MCI: Sensitivität 80 %, Spezifität 92 %. Für MCI und leichte Demenz: Sensitivität 100%, Spezifität 92%. Für vaskuläre Demenz: Sensitivität und Spezifität je über 95%.

Brand (2002): n = 231, gesunde Probanden aus drei Altersgruppen (bis 30, 45-60, über 60 J). Rohwerte nicht (berufs-)bildungs-, aber altersabhängig. Nach Transformation der Rohwerte gemäß Erstautoren ist der Alterseffekt aufgehoben.

Kessler (2000): n = 175 gesunde Personen, davon 93 über 60 J 169 mit (V. a.) Alzheimer Demenz bis zu mittlerer Schwere. Sensitivität für Alzheimer-DemenzEs gibt eine validierte türkische Version (Aycicek et al. 2020).

Begründung der Aufnahme in die Leitlinie:

Sensitiver als MMSE in der Früherkennung einer demenziellen Entwicklung.

Limitationen / Forschungsbedarf:

Anforderungen an das Hörvermögen, das Sehvermögen (Lesen) und die Feinmotorik (Schreiben). Proband*innen, in deren Leben der Umgang mit Zahlen keine Rolle spielte, können durch Punktverlust beim Zahlenumwandeln leicht fehleingeschätzt werden!

Literatur:

Aycicek GS, Çalışkan H, Ozsurekci Cv et al. (2020). A Reliable Tool for Assessing MCI and Dementia: Validation Study of DemTect for Turkish Population. American Journal of Alzheimer's Disease & Other Dementias 35. DOI:10.1177/1533317520949805

Brand M, Kalbe E, Kessler J et al. (2002) Test performance in dementia screenings are influenced by age but not by education and profession – new data of the DemTect. Neurorehabilitation and Neural Repair. 15:301

Kalbe E, Kessler J, Calabrese P et al. (2004) DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. Int J Geriatr Psychiatry 19:136-143

Kalbe E, Brand M, Kessler J, Calabrese P (2005) Der DemTect in der klinischen Anwendung. Zeitschrift für Gerontopsychologie & -psychiatrie 18:121–130 DOI 10.1024/1011-6877.18.3.121

Kessler J, Calabrese P, Kalbe E, Berger F (2000) DemTect: Ein neues Screening-Verfahren zur Unterstützung der Demenzdiagnostik. Psycho 26:343–347

Kohn N, Kalbe E, Georg H, Kessler J (2007) Vergleich MMST und DemTect: Spezifität und Sensitivität bei primär kognitiven Störungen. Aktuelle Neurologie, 34 – P672. DOI: 10.1055/s-2007-987943

Ozer S, Young J, Champ C et al. (2016) A systematic review of the diagnostic test accuracy of brief cognitive tests to detect amnesic mild cognitive impairment. International Journal of Geriatric Psychiatry. 31: 1139-1150

Zhuang L, Yang Y, Gao J (2021) J Neurol 268(5):1615-1622. DOI: 10.1007/s00415-019-09506-7

Tab. 2.4.13: Kognition

Name des Instrumentes:	Montreal Cognitive Assessment		
Kürzel:	MoCA		
Mittlerer Zeitaufwand:			
Für Patient*in 12 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 14 min			
Lernaufwand für Untersucher*in:			
O leicht	x mittel	O anspruchsvoll	
Durchführung durch:			
Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
x leicht/Früherkennung	x mittel	O schwer	
Gütekriterien:			
Bernini (2025): Studie aus Italien; 1203 gesunde Personen, 93 Pat. mit Alzheimer-Demenz und 127 Pat. mit Parkinson-Demenz wurden mit MMSE und MoCA untersucht, Konversionstabellen zur Umwandlung von MMSE-Scores zu MoCA Scores und vice versa vorgestellt. Die Verlässlichkeit der Umwandlung ist bei höheren Testwerten besser. Stewart (2022): Zusammenfassung von relevanten Cochrane-Reviews. Die Autor*innen schlussfolgern, in der Primärversorgung erreiche das MoCA keine ausreichende prädiktive Validität im Hinblick auf eine Demenzentwicklung Davis (2021), Cochrane Review: Bei Cut-Off von 27/26 über die eingeschlossenen Studien hinweg hohe Sensitivität von 94% oder mehr, aber geringe Spezifität von 60% oder weniger. Pinto (2019): In einem systematischen Review kommen die Autor*innen zu dem Schluss, dass MMSE und MoCA zur Demenzdetektion gleich geeignet sind, MoCA der MMSE zur Diagnose eines MCI aber überlegen ist. Carson (2018): Bei Cut-off 27/26 Sensitivität 94%, Spezifität 66%, bei Cut-off 24/23 Sensitivität 83%, Spezifität 88%, höherer Anteil korrekt klassifizierter Patient*innen. Thomann (2018): n = 283 aus einem Register potenziell an Studien Teilnehmender rekrutierte, kognitiv gesunde Personen (CERAD-NAB, Mini Mental State Examination) ohne Depression, 73,8 ± 5,2 (65 – 91) J, 54,8% F. Korrelation mit CERAD-NAB 0,45, mit Mini Mental State Examination 0,20. Höhere Bildung war stärker als niedrigeres Alter, dieses stärker als weibliches Geschlecht mit höheren Werten assoziiert. 31,1% der Proband*innen erreichten unter 26 Punkte. Tafeln weisen u. a. die zu einer Spezifität von 90% passenden bildungs-, alters-und geschlechtsadaptierten Werte aus (s. u.). Ciesielska (2016), Metaanalyse: Bei Cut-off 25/24 zur Erkennung von MCI Sensitivität 80%, Spezifität 81%. Davis (2015), Cochrane Review: Bei Cut-off 19/18 Sensitivität von 86%, Spezifität von 97%, bei Cut-off 27/26 Sensitivität 100%, Spezifität 34%. Nasreddine (2005): Kontrollgruppe 72,84 ± 7,03 J; Gruppe MCI 75,19 ± 6,27 J; Gruppe Alzheimer-Demenz 76,72 ± 8,83 J. Sensitivität von 90% für MCI und 100% für Alzheimer-			

Demenz bei Cut-off von 26/25.

Begründung der Aufnahme in die Leitlinie:

Sensitiver als MMSE in der Früherkennung einer demenziellen Entwicklung.

Limitationen / Forschungsbedarf:

Keine Stratifizierung nach Alter/Ausbildung/Geschlecht. Anforderungen an das Hörvermögen (siehe MoCA-HI), das Sehvermögen (Lesen, Zeichnen) und die Feinmotorik (Schreiben, Zeichnen).

Literatur:

<https://mocatest.maskin.ch/de/standardwerte/normwerttabellen-herunterladen> Bernini S, Ballante E, Picascia M et al. (2025) Equating conversion norms for the Mini-Mental State Examination and Montreal Cognitive Assessment in healthy subjects and patients with neurodegenerative disorders. *Exp Gerontol* 205:112756 DOI: 10.1016/j.exger.2025.112756

Carson N, Leach L, Murphy KJ (2018) A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *Int J Geriatr Psychiatry* 33:379-388

Ciesielska N, Sokołowski R, Mazur E. et al. (2016) Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60? Meta-analysis. *Psychiatr Pol* 50:1039-1052

Davis DHJ, Creavin ST, Yip JLY et al. (2021) Montreal Cognitive Assessment for the detection of dementia. *Cochrane Database of Systematic Reviews* 2021, Issue 7. Art. No.: CD010775. DOI: 10.1002/14651858.CD010775.pub3.

Davis DH, Creavin ST, Yip JL et al. (2015) Montreal Cognitive Assessment for the diagnosis of Alzheimer's disease and other dementias. *Cochrane Database of Systematic Reviews* 2015, Issue 10. Art. No.: CD010775. DOI: 10.1002/14651858.CD010775.pub2.

Horstmann S, Rizos T, Rauch G et al. (2014) Feasibility of the Montreal Cognitive Assessment in acute stroke patients. *Eur J Neurol* 21:1387-1393 DOI: 10.1111/ene.12505

Mai LM, Sposato LA, Rothwell PM. Et al. (2016) A comparison between the MoCA and the MMSE visuoexecutive sub-tests in detecting abnormalities in TIA/stroke patients. *Int J Stroke* 11:420-424

Nasreddine, ZS, Phillips NA, Bedirian V. et al. (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53:695-699 doi: 10.1111/j.1532-5415.2005.53221.x. Erratum in: *J Am Geriatr Soc.* (2019) 67(9):1991. doi: 10.1111/jgs.15925

Pinto T, Machado L, Bulgacov T et al. (2019) Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's Disease (AD) in the elderly? *International Psychogeriatrics* 31: 491-504

Roalf DR, Moberg PJ, Xie SX et al. (2013) Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimers Dement* 9:529-537

Saczynski JS, Inouye SK, Guess J et al. (2015) The Montreal Cognitive Assessment: Creating a Crosswalk with the Mini-Mental State Examination. *J Am Geriatr Soc* 63:2370-2374

Scheffels JF, Kräling H, Kalbe E, Kessler J (2018) Konversionen von kognitiven Screenings. Mini-Mental-Status-Test vs. Montreal Cognitive Assessment vs. DemTect. *Nervenarzt* 89:1371-1377

Shi D, Chen X, Li Z (2018) Diagnostic test accuracy of the Montreal Cognitive Assessment in the detection of post-stroke cognitive impairment under different stages and cutoffs: a systematic review and meta-analysis. *Neurol Sci* 39:705-716

Stewart C, Stewart J, Stubbs S et al. (2022) Mini-Cog, IQCODE, MoCA, and MMSE for the Prediction of Dementia in Primary Care. *Am Fam Physician* 105: 590-592

van Steenoven I, Aarsland D, Hurtig H et al. (2014) Conversion between mini-mental state examination, Montreal Cognitive Assessment, and dementia rating scale-2 scores in Parkinson's

disease. Mov Disord 29:1809-1815

Thomann AE, Goettel N, Monsch RJ et al. (2018) The Montreal Cognitive Assessment: Normative Data from a German-Speaking Cohort and Comparison with International Normative Samples. Journal of Alzheimer's Disease 64:643–655. DOI 10.3233/JAD-180080

Trzepacz PT, Hochstetler H, Wang S. et al. (2015) Alzheimer's Disease Neuroimaging Initiative. Relationship between the Montreal Cognitive Assessment and Mini-mental State Examination for assessment of mild cognitive impairment in older adults. BMC Geriatr 15:107

Wang G, Estrella A, Hakim O, Milazzo P et al. (2022) Mini-Mental State Examination and Montreal Cognitive Assessment as Tools for Following Cognitive Changes in Alzheimer's Disease Neuroimaging Initiative Participants. J Alzheimers Dis 90(1):263-270

Wong A, Black SE, Yiu SY et al. (2018) Converting MMSE to MoCA and MoCA 5-minute protocol in an educationally heterogeneous sample with stroke or transient ischemic attack. Int J Geriatr Psychiatry 33:729-734

Literatur zu MoCA-Kurzformen:

Cecato JF, Martinelli JE, Izbicki R et al. (2016) A subtest analysis of the Montreal Cognitive Assessment (MoCA): which subtests can best discriminate between healthy controls, mild cognitive impairment and Alzheimer's disease? Int Psychogeriatr 28:825-832

Horton DK, Hynan LS, Lacritz LH et al. (2015) An abbreviated Montreal Cognitive Assessment (MoCA) for dementia screening. Clin Neuropsychol 29:413-425

Roalf DR, Moore TM, Mechanic-Hamilton D et al. (2017) Bridging cognitive screening tests in neurologic disorders: A crosswalk between the short Montreal Cognitive Assessment and Mini-Mental State Examination. Alzheimers Dement 13:947-952

Tabelle 2.4.14 Kognition

Name des Instrumentes:	Montreal Cognitive Assessm.-Hearing Impairment	
Kürzel:	MoCA-HI	
Mittlerer Zeitaufwand:		
Für Patient*in ca. 20 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 25 min		
Lernaufwand für Untersucher*in:		
☐ leicht	☒ mittel	☐ anspruchsvoll
Durchführung durch:		
Arzt/Ärztin, Psycholog*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien:		
Völter (2023): Studienpopulation n = 346, mittleres Alter 77,2 J (60-97 J), 57,8 % F. Spezifität 90 % bei Cut-off 22/21. 19-21 Punkte sprechen für eine MCI <19 für eine Demenz. Bei Cut-off 25/24 Sensitivität 97,5 % für die Erfassung von MCI. Im Graubereich von 22-24 Punkten ist eine ausführliche neuropsychologische Testung empfohlen sowie eine Test-Wiederholung nach 6-12 Monaten.		

Dawes (2023): Studienpopulation n = 256 (Gruppe 1 mit Höreinschränkung n = 159, Gruppe 2 ohne Höreinschränkung n = 97, bei n=28 Retestung nach 2-4 Wochen), Alter ab 60 J, mittleres Alter 74,8 J, 51,2 % F. AUC 0,973. Bei Cut-off 25/24 Sensitivität von 92,8 % und Spezifität von 90,8 % für das Erkennen von Demenz. Retest-Reliabilität 0,92 (95%-KI 0,78 - 0,97). Interne Konsistenz Cronbachs α 0,91.

Völter (2022): Studienpopulation n=175 (Gruppe 1 ohne/mit nur leichter Höreinschränkung n=94, Gruppe 2 mit mittlerer/starker Höreinschränkung n=81), mittleres Alter 71,5 J (60-97 J), 43,4 % F. MoCA-HI-Ergebnisse ohne signifikanten Gruppenunterschied (p=0,05). Beim Standard-MoCA Gruppe 1 signifikant besser als Gruppe 2 (p<0,001). Kein signifikanter Unterschied zwischen MoCA- und MoCA-HI-Leistung in Gruppe 1 (p=0,12), aber Gruppe 2 beim MoCA-HI signifikant besser als beim Standard-MoCA (p<0,001). Retest-Reliabilität des MoCA-HI 0,84.

Begründung der Aufnahme in die Leitlinie:

Ermöglicht die Nutzung des zur MCI-Erfassung geeigneten MoCA auch bei schwerer Hörstörung, einer im Alter häufigen Beeinträchtigung.

Limitationen / Forschungsbedarf:

Der Test MoCA-HI ist bisher nur in der Version 8.1 verfügbar, so dass Re-Testungen mit anderen Versionen derzeit nicht möglich sind.

Literatur:

Verfügbar unter <https://mocacognition.com/paper/> (Zugriff 14.11.2024)

Dawes P, Pye A, Reeves D et al. (2019) Protocol for the development of versions of the Montreal Cognitive Assessment (MoCA) for people with hearing or vision impairment. BMJ Open 9(3):e026246. doi: 10.1136/bmjopen-2018-02624

Dawes P, Reeves D, Yeung WK et al. (2023) Development and validation of the Montreal cognitive assessment for people with hearing impairment (MoCA-H). J Am Geriatr Soc. 2023 May;71(5):1485-1494. doi: 10.1111/jgs.18241. Epub 2023 Feb 1. PMID: 36722180.

Theocharous S, Savage G, Charalambous AP et al. (2024) A cross-cultural study of the Montreal Cognitive Assessment for people with hearing impairment. J Am Geriatr Soc 72(10):3156-3162. doi: 10.1111/jgs.19020

Völter C, Fricke H, Götze L et al. (2022) Evaluation of the non-auditory neurocognitive test MoCA-HI for hearing-impaired. Front Neurol 13:1022292. doi: 10.3389/fneur.2022.1022292

Völter C, Fricke H, Faour S et al. (2023) Validation of the German Montreal-Cognitive-Assessment-H for hearing-impaired. Front Aging Neurosci 15:1209385. doi: 10.3389/fnagi.2023.1209385.

Tab. 2.4.15: Kognition

Name des Instrumentes:	Test zur Früherkennung von Demenzen mit Depressionsabgrenzung
Kürzel:	TFDD (Englisch: TE4D)
Mittlerer Zeitaufwand: Für Patient*in 12 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 15 min	

Lernaufwand für Untersucher*in:		
O leicht	x mittel	O anspruchsvoll
Durchführung durch: Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
x leicht /Früherkennung	O mittel	O schwer
Gütekriterien: Ihl (2000): Altersmedian Gruppe Alzheimer-Demenz 81 J, Gruppe depressive Störung 71 J, gesunde Kontrollgruppe 56. Sensitivität 98,4%, Spezifität 97,2% bei Patient*innen ohne und 96,9% gegenüber Patient*innen mit depressiver Störung.		
Begründung der Aufnahme in die Leitlinie: Sensitiver als MMSE in der Früherkennung einer demenziellen Entwicklung.		
Limitationen / Forschungsbedarf: Keine Stratifizierung nach Alter/Ausbildung/Geschlecht. Anforderungen an das Hörvermögen, Sehvermögen (Lesen) und die Feinmotorik (Schreiben). Weitere Studien liegen nicht vor.		
Literatur: Brinkmeyer J, Grass-Kapanke B, Ihl R (2004) EEG and the Test for the Early Detection of Dementia with Discrimination from Depression (TE4D): a validation study. Int J Geriatr Psychiatry 19(8):749-753. Doi: 10.1002/gps.1162 Grass-Kapanke B, Brieber S, Pentzek M, Ihl R (2005) Der TFDD Test zur Früherkennung von Demenzen mit Depressionsabgrenzung. Zeitschrift für Gerontopsychologie & -psychiatrie 18:155–167 Ihl R, Grass-Kapanke B, Lahrem P et al. (2000) Entwicklung und Validierung eines Tests zur Früherkennung der Demenz mit Depressionsabgrenzung (TFDD). Fortschr Neurol Psychiatr 68:413-422 Mahoney R, Johnston K, Katona C et al. (2005) The TE4D-Cog: a new test for detecting early dementia in English-speaking populations. Int J Geriatr Psychiatry 20(12):1172-1179. DOI: 10.1002/gps.1412		

Tab. 2.4.16: Kognition

Name des Instrumentes:	Global Deterioration Scale (Reisberg-Skala)	
Kürzel:	GDS (kog)	
Mittlerer Zeitaufwand: Entfällt für Patient*in, für Untersucher*in (inkl. Vor- u. Nachbereitung) 20 min		
Lernaufwand für Untersucher*in:		
☐ leicht	☐ mittel	☒ anspruchsvoll

Durchführung durch:

Ärzt*in, Psycholog*in

Geeignet für Schweregrad:

x leicht/Früherkennung	x mittel	x schwer
-------------------------------	-----------------	-----------------

Gütekriterien:

Reisberg (1982): 38 Patient*innen, MW 69,3 ± 7,9 Jahre; 58% Frauen

→ Interrater-Reliabilität 0,97; Retest-Reliabilität 0,92 (Korrelation in einem 7-Tage- bis 4-Monate-Intervall)

Korrelation der GDS (kog) mit 25 von 26 psychometrischen Assessment-Instrumenten signifikant ($r = 0,3-0,6$, $p < 0,05$).

Im Gerontopsychiatrischen Zentrum des PSD Wien konnten bei 4475 Erstkontakt-Patient*innen (MW 76 Jahre; 67 % Frauen; ambulantes Setting) folgende Korrelationen erhoben werden: Interrater-Reliabilität 0,97; GDS (kog) / MMSE: -0,92; GDS (kog) / TFDD: -0,89

Begründung der Aufnahme in die Leitlinie:

Fremdbeurteilungsinstrument, das den Einbezug anamnestischer Angaben Angehöriger erlaubt und über das Vorhandensein typischer Symptome eine Stadieneinteilung der Alzheimer-Demenz ermöglicht.

Limitationen / Forschungsbedarf:

In der Reisberg-Skala wird der typische Verlauf der degenerativen Demenz vom Alzheimerstyp abgebildet, bei anderen Demenzformen, wie vaskulären oder frontalen Demenzerkrankungen, ist die Skala nicht passend.

Literatur:

Reisberg B, Ferris SH, de Leon MJ, Crook T (1982) The global deterioration scale for assessment of primary degenerative disease. Am J Psychiatry 139: 1136-1139

Reisberg B, Ferris SH, de Leon MJ, Crook T (1988) Global Deterioration Scale (GDS). Psychopharmacol Bull 24:661-663

Reisberg B, Torossian C, Shulman MB et al. (2019) Two Year Outcomes, Cognitive and Behavioral Markers of Decline in Healthy, Cognitively Normal Older Persons with Global Deterioration Scale Stage 2 (Subjective Cognitive Decline with Impairment). J Alzheimers Dis 67(2):685-705

Rodríguez-Hidalgo E, García-Alba J, Novell R, Esteba-Castillo S (2023) The Global Deterioration Scale for Down Syndrome Population (GDS-DS): A Rating Scale to Assess the Progression of Alzheimer's Disease. Int J Environ Res Public Health 20(6):5096. <https://doi.org/10.3390/ijerph20065096>

Tab. 2.4.17: Vergleich der Scores des DemTect mit denen der MMSE und des MoCA (Scheffels 2018)

DemTect	3	6	9	12	15	18
MMSE	17	22	26	28	29	30
MoCA	9	15	20	24	27	30

Tab. 2.4.18: Vergleich der Scores des MoCA mit denen der MMSE

MoCA	5	10	15	20	23	24	25	26	27	28	
MMSE	11	16	22	26	28	28	29	29	29	30	Roalf et al. 2013
MMSE	12	18	21	25	27	28	29	29	30	30	Saczynski et al. 2015
MMSE	9	18	22	25	27	27	28	28	29	29	Scheffels et al. 2018
MMSE	13	17	21	25	27	28	28	29	29	30	van Steenoven et al. 2014
MMSE	11	17	23	26		28	29			30	Wong et al. 2018

Beispiele:

- 1. Patient*in anamnestisch und klinisch ohne Hinweise auf kognitive Beeinträchtigung.
SIS 6 Punkte (Stufe 2a)
→**Weiteres Assessment der Kognition zurzeit nicht medizinisch indiziert.**
- 2. Patient*in mit mehrjährigem Demenzverlauf, keine Testwerte bekannt.
MMSE von Patient*in nach 12 Items (bis dahin 2 Punkte) abgebrochen bei massiver Überforderung, klinisch und ggf. CAM/DOSS/NuDesc kein Hinweis auf Delir
→ **im Verlauf SMMSE**
- 3. Patient*in mit schwerem Visusverlust. Angehörige vermuten eine demenzielle Entwicklung.
MMSE-blind 16 P (Hinweis auf mögliches kognitives Defizit)
→ **BAS 20 Punkte (passend zu Demenz vom Alzheimer-Typ)**
→ **i. d. R. weitere Diagnostik (Labor, MRT)**
- 4. Patient*in mit derzeit kontinuierlichem Sauerstoff-Bedarf bei infektexazerbierter COPD und Dyspnoe bei niedriger Belastung. Klinisch kein Hinweis auf kognitive Störung, Patient*in gibt jedoch an, abnehmende Gedächtnisleistungen festzustellen.
SIS 5 Punkte am Aufnahmetag
→**nach einer Woche bei besserer Belastbarkeit (und Konzentrationsfähigkeit)**
MoCA/DemTect/TFDD: kein Hinweis auf kognitive Störung
→**Depression ausschließen, Kontrolle Kognition in sechs Monaten (bei**

entsprechenden Hinweisen früher)

5. Patient*in mit deutlich pathologischem hausärztlichen kognitiven Assessment (z. B. **Uhrentest/MiniCog/3-Wörter-Uhrentest**), Labor ohne Hinweise zu möglichen Ursachen, MRT ist angemeldet, noch keine weitere neuropsychologische Abklärung erfolgt.
→sofern innerhalb der ersten Behandlungstage verfügbar direkter Übergang auf Stufe 3 des Assessments (z. B. CERAD-Plus)

Delirerfassung

Der Einsatz der hier genannten Instrumente kann zum früheren Erkennen eines Delirs beitragen.

Mit der **Nursing Delirium Screening Scale (NU-DESC)** (Tabelle 2.4.19) wurde ein pflegebasiertes, DSM-IV adaptiertes Instrument zur Delirdetektion entwickelt, welches aus der Confusing Rating Scale (CRS) hervorgegangen ist und um das Kriterium der „psychomotorischen Retardierung“ erweitert wurde. Dadurch sollte dem Umstand des oft unerkannten hohen Anteils hypoaktiver Delirien Rechnung getragen werden. Zudem konnte gezeigt werden, dass durch den Einsatz dieses Instrumentes nicht nur signifikant mehr Delirien erkannt werden, sondern hierdurch die Detektion auch zu einem früheren Zeitpunkt gelingt. Auf der Normalstation erwies sich die Nu-DESC gegenüber der Delirium Observation Screening Scale (DOSS, s. u.) als äquipotent hinsichtlich Sensitivität, Spezifität, positivem und negativem Vorhersagewert und kann dabei gut in den Routine-Alltag implementiert werden. Das Instrument ist besonders gut geeignet für chirurgische Settings. In einem „head-to-head“-Vergleich mit dem 4AT (s. u.) schnitt die Nu-DESC bei den Pflegekräften hinsichtlich Akzeptanz, Praktikabilität und Integration in die Alltagsroutine deutlich besser ab. Die aus fünf Items bestehende NU-DESC wurde in mehrere Sprachen aus dem Englischen übersetzt, und es liegt auch eine validierte deutschsprachige Version vor. Eine Metaanalyse [Lin 2025] vergleicht die diagnostische Genauigkeit von Nu-DESC und Confusion Assessment Method (CAM, s. u.) beim postoperativen Delir. Die Ergebnisse zeigen eine vergleichbare Sensitivität, jedoch eine kürzere Durchführungszeit für Nu-DESC. Es wird festgehalten, dass die Sensitivität von Nu-DESC (gepoolte Sensitivität 0,69) und CAM (gepoolte Sensitivität 0,65) zur Erkennung eines postoperativen Delirs kein optimales Niveau erreichen.

Der **4 „A“s Test (4AT)** (Tabelle 2.4.20) erfordert keine spezielle Schulung. Die vier „A“s stehen für **A**lertness (Wachheit), **A**bbreviated Mental Test- 4 (Orientierung - korrektes Nennen von Alter, Geburtsdatum, aktuellem Ort, aktuellem Kalenderjahr), **A**ttention (Aufmerksamkeit - Monate rückwärts aufzählen) und **A**cute change (akute oder fluktuierende Symptomatik). Auch bei Patienten, die für kognitive Tests oder Befragungen zu somnolent oder unruhig sind, gibt es einen klinisch verwertbaren 4AT-Score. Der 4AT ist zurzeit das in der Literatur am meisten validierte Instrument zur Erkennung von Delirien, er wurde in der Akutmedizin, Chirurgie, Notaufnahme, stationären Pflege, in Hospizen und Schlaganfall-Stationen validiert. Sensitivität und Spezifität sind in unterschiedlichen Settings sehr gut. Der 4AT ist heute eines der weltweit am häufigsten verwendeten Instrumente, mit hohen Ausfüllraten und einer exzellenten

Delir-Erkennung. Er behält auch bei wiederholter Durchführung seine diagnostische Genauigkeit und ist für die Beurteilung des Delir-Verlaufs geeignet.

Die **Delirium Observation Screening Scale (DOSS)** (Tabelle 2.4.21) wurde zur raschen systematischen Erfassung und Dokumentation von Auffälligkeiten der Kognition und des Verhaltens, die auf ein Delir hinweisen könnten, entwickelt. Die 13 Fragen, die sich auf Beobachtungen im Rahmen der üblichen Pfl egetätigkeiten beziehen, sind jeweils von Tag- und Nachtdienst zu beantworten. Die Summe der Auffälligkeiten wird addiert und das arithmetische Mittel über die Beobachtungen in jeweils 24 Stunden gebildet. Es werden sowohl Charakteristika eines hyper- als auch eines hypoaktiven Delirs erfasst. Bei einem Wert von ≥ 3 besteht der Verdacht auf das Vorliegen eines deliranten Zustandsbildes.

Die korrekte Anwendung der **Confusion Assessment Method (CAM)**, auch CAM-long) (Tabelle 2.4.22) erfordert ein spezifisches Training der durchführenden Personen, um das Interview korrekt durchzuführen und zu bewerten, da es sonst zu einer hohen Fehlerquote kommt. Es besteht ein Copyright (Hospital Elder Life Program), die American Geriatric Society erteilt die Nutzungserlaubnis. Die American Delirium Society führt als „Mitglieder der CAM-Familie“ mittlerweile folgende Varianten auf [<https://americandeliriumsociety.org/healthcare-professionals/ags-cocare-cam-and-help-tools/>]: CAM-long (10 Items), CAM-short (deren erste 4 Items; für die Notaufnahme geeignet), CAM-S (Ergänzung zur CAM-long, um den Schweregrad zu erfassen), FAM-CAM (Interview informell Pfl egender), 3D-CAM (Patient*innen-Interview mit 20 plus 2 optionalen Fragen, für das eine Bearbeitungsdauer von drei Minuten angegeben wird, erweiterbar um eine Schweregrad-Erfassung) und UB-CAM (zweischrittiges adaptives Testen unter Einsatz von Items der 3D-CAM, Testdauer im Mittel eine Minute).

Tab. 2.4.19: Delirerfassung

Name des Instrumentes:	Nursing Delirium Screening Scale		
Kürzel:	NU-DESC		
Mittlerer Zeitaufwand:			
Entfällt für Patient*in / für Untersucher*in (inkl. Vor- u. Nachbereitung) 1 min			
Lernaufwand für Untersucher*in:			
x leicht		O mittel	O anspruchsvoll
Durchführung durch:			
Pflegekraft			
Geeignet für Schweregrad:			
x leicht/Früherkennung		x mittel	x schwer
Gütekriterien:			
In der Validierungsstudie wurden gute Testgütekriterien bei einer Sensitivität von 86% und			

einer Spezifität von 87% ermittelt. Im Vergleich mit der CAM zeigt die NU-DESC eine gute Leistungsfähigkeit bei einem positiven prädiktiven Wert von 0,77 und einem negativen prädiktiven Wert von 0,91. Im Setting einer chirurgischen Station zeigte die NU-DESC noch bessere Testgütekriterien mit einer Sensitivität von 98% und einer Spezifität von 92%.

De la Cruz (2015): RCT bei Krebspatienten in der letzten Lebenswoche im Hospiz, n = 78, 69 (49-91) J, 55% F. Bei Cut-off 1/2 Sensitivität 63%, Spezifität 67%, positiver prädiktiver Wert 0,61.

Begründung der Aufnahme in die Leitlinie:

Mit wenig Aufwand in die pflegerische Routine integrierbares Instrument.

Limitationen / Forschungsbedarf:

Manche Autoren schlagen zur Verbesserung der Sensitivität einen Cut-off von 0/1 (statt 1/2) vor oder empfehlen einen zusätzlichen Aufmerksamkeitstest, um ohne nennenswerten Verlust der Spezifität die Sensitivität zu erhöhen. Eine andere Arbeitsgruppe empfiehlt die Operationalisierung des Nu-DESC-Items „Desorientierung“ durch eine zusätzliche Abfrage des Wochentages und des Namens der behandelnden Krankenhausfachabteilung, wodurch gleichfalls die Sensitivität signifikant erhöht werden könne.r

Literatur:

Bergjan M, Zilezinski M, Schwalbach T, Franke C, Erdur H, Audebert HJ, Hauß A. (2020) Validation of two nurse-based screening tools for delirium in elderly patients in general medical wards. BMC Nurs. 19:72 DOI: 10.1186/s12912-020-00464-4

Brich J, Baten V, Wußmann J, Heupel-Reuter M, Perlov E, Klöppel S, Busch HJ. (2019) Detecting delirium in elderly medical emergency patients: validation and subsequent modification of the German Nursing Delirium Screening Scale. Intern Emerg Med. 14(5):767-776

De la Cruz M, Noguera A, San Miguel-Arregui MT et al. (2015) Delirium, Agitation, and Symptom Distress within the Final Seven Days of Life among Cancer Patients Receiving Hospice Care. Palliat. Support. Care 13:211–216. DOI: 10.1017/S1478951513001144

Gaudreau JD, Gagnon P, Harel F et al. (2005) Fast, systematic, and continuous delirium assessment in hospitalized patients: the nursing delirium screening scale. J Pain Symptom Manage 29:368-375

Heinrich TW, Kato H, Emanuel C, Denson S. (2019) Improving the Validity of Nurse-Based Delirium Screening: A Head-to-Head Comparison of Nursing Delirium-Screening Scale and Short Confusion Assessment Method. Psychosomatics. 60(2):172-178

Lin CJ, Fick DM, Traynor V et al. (2025) Comparative Diagnostic Accuracy of Nursing Delirium Screening Scale Versus Confusion Assessment Method for Postoperative Delirium: A Systematic Review and Meta-Analysis. J clin nurs 34(1):287–298

Radtke FM, Franck M, Schust S et al. (2010) A comparison of three scores to screen for delirium on the surgical ward. World J Surg 34:487-494

Zastrow I, Tohsche P, Loewen T, Vogt B, Feige M, Behnke M, Wolff A, Kiefmann R, Olotu C. (2021) Comparison of the '4-item assessment test' and 'nursing delirium screening scale' delirium screening tools on nonintensive care unit wards: A prospective mixed-method approach. Eur J Anaesthesiol. 38(9):957-965. DOI: 10.1097/EJA.0000000000001470

Tab. 2.4.20: Delirerfassung

Name des Instrumentes: Kürzel	4 „A“s Test 4AT		
Mittlerer Zeitaufwand: Für Patient*in 2 min, für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch: Pflegekraft, Ärzt*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien: Lin (2023): Der 4AT wurde im Vergleich zur CAM gegen etablierte Diagnosekriterien (ICD - International Statistical Classification of Diseases and Related Health Problems; DSM – Diagnostic and Statistical Manual of Mental Disorders) evaluiert. Gepoolte Sensitivität 76% versus 80%, gepoolte Spezifität 90% versus 99%. Tieges (2021): systematisches Review mit Metaanalyse, umfasste 17 Studien: Gepoolte Sensitivität 88% (95% CI 80-93%), gepoolte Spezifität 88% (95% CI 82-92%). Jeong (2020): systematisches Review mit Metaanalyse aus 13 Studien: Gepoolte Sensitivität 81,5 % (95% CI 70,7-89,0%), gepoolte Spezifität. 87,5 % (79,5-92,7%). Bellelli (2014): Validierungsstudie bei älteren stationären Krankenhauspatient*innen, Sensitivität 90% und Spezifität 84%. Begründung der Aufnahme in die Leitlinie: In verschiedenen Settings validiertes, international angewandtes, leicht zu erlernendes Instrument zur Erkennung von Delirien mit hoher diagnostischer Zuverlässigkeit. Limitationen / Forschungsbedarf: Keine Studien zur Anwendung in der Intensivmedizin. Literatur: Bellelli G, Morandi A, Davis DH et al. (2014) Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. Age Ageing 43(4):496-502. DOI: 10.1093/ageing/afu021. Erratum in: Age Ageing 2015 44(1):175 Jeong E, Park J, Lee J (2020) Diagnostic Test Accuracy of the 4AT for Delirium Detection: A Systematic Review and Meta-Analysis. Int J Environ Res Public Health 17(20):7515. DOI: 10.3390/ijerph17207515 Lin CJ, Su IC, Huang SW (2023) Delirium assessment tools among hospitalized older adults: A systematic review and metaanalysis of diagnostic accuracy. Ageing Res Rev 90:102025. DOI: 10.1016/j.arr.2023.102025 MacLulich A, Ryan T, Cash H (2011) The 4A Test: Screening instrument for cognitive impairment and delirium. Edinburgh, Scotland: University of Edinburgh and NHS Lothian. https://www.the4at.com/4at-deutsche (Zugriff 14.11.2024)			

MacLulich AM, Shenkin SD, Goodacre S et al. (2019) The 4 'A's test for detecting delirium in acute medical patients: a diagnostic accuracy study. Health Technol Assess 23(40):1-194. DOI: 10.3310/hta2340

Tieges Z, MacLulich AMJ, Anand A et al. (2021) Diagnostic accuracy of the 4AT for delirium detection in older adults: systematic review and meta-analysis. Age Ageing 50(3):733-743. DOI: 10.1093/ageing/afaa224

Tabelle 2.4.21: Delirerfassung

Name des Instrumentes:	Delirium Observation Screening Scale	
Kürzel:	DOSS	
Mittlerer Zeitaufwand:		
Entfällt für Patient*in / für Untersucher*in (inkl. Vor- u. Nachbereitung) 5 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Pflegefachkraft		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien:		
Martinez-Arnau (2023): Systematisches Review bezüglich der Gütekriterien von 13 Instrumenten zur Delirerfassung bei Krebspatient*innen. Keine relevante Studie hatte ein höheres Alter als Einschlusskriterium definiert, 14 mit einem mittleren Alter der Teilnehmenden >60 J wurden aufgenommen, davon fokussierten 2 die DOSS. Die Autor*innen favorisieren diese als das Instrument mit den besten Gütekriterien. Die DOSS wurde in 2 Hochrisikogruppen (Geriatrische Spitalspatient*innen und Hüftfrakturpatient*innen) entwickelt und in der Folge in verschiedenen Settings getestet: Validität: r = 0,63 gegenüber der CAM Sensitivität: 100% bei Patient*innen nach Herzoperation, 97% bei zuhause betreuten Hospizpatient*innen, 90% bei allg. internistischen Spitalspatient*innen ab 65 Jahren (Iowa), 82% bei Palliativpatient*innen im Spital. Spezifität: 96% bei Patient*innen nach kardiochirurgischem Eingriff. 96% bei Palliativpatient*innen im Spital (positiver prädiktiver Wert 0,69, negativer prädiktiver Wert 0,98), 91% bei allg. internistischen Spitalspatient*innen ab 65 Jahren (Iowa), 89% bei zuhause betreuten Hospizpatient*innen. Für Pflegeheimpatient*innen (ca. 50% an Demenz erkrankt) ergab sich trotz Hinaufsetzens der zutreffenden Kriterien von 3 auf 5 in 24 Std. eine niedrige Sensitivität (61%) und Spezifität (71%). Eine rezente Metaanalyse bestätigte die hohe Sensitivität (90%; 76%–97%, CI 95%) und Spezifität (92%; 88%–94%, CI 95%).		

Neefjes (2019): Stationäre Patient*innen mit fortgeschrittener onkologischer Erkrankung, Delirium Rating Scale (DRS-R-98) als Goldstandard, n = 193, Alter bei den 93 positiv getesteten Patient*innen im Mittel 68/bei den negativ getesteten 60 J, 39%/28% F. Bei Cut-off von 2/3 Sensitivität der DOSS fast 100% (95%-Konfidenzintervall 95,8-100,0%), Spezifität 99,5%, positiver prädiktiver Wert 94,6%, negativer prädiktiver Wert fast 100%.

Begründung der Aufnahme in die Leitlinie:

Laut aktuellem Review (s. o.) zuverlässigstes im Rahmen der Pflege einzusetzendes Instrument zur Delir-Erkennung.

Limitationen / Forschungsbedarf:

Die Diagnosesicherung erfordert ein weiteres Instrument, z. B. Confusion Assessment Method oder Delirium Rating Scale.

Literatur:

De J, Wand AP (2015) Delirium Screening: A Systematic Review of Delirium Screening Tools in Hospitalized Patients. *Gerontologist*. 55(6):1079-1099. doi: 10.1093/geront/gnv100. Erratum in: *Gerontologist*. 2017 57(2):387. doi: 10.1093/geront/gnw120

Detroyer E., Clement PM, Milisen K et al. (2014) Detection of delirium in palliative care unit patients: a prospective descriptive study if the Delirium Observation Scale administered by bedside nurses. *Palliat.Med* 28:79-86

Gavinski K, Carnahan R, Weckmann M (2016) Validation of the Delirium Observation Screening Scale in a hospitalized older population. *J Hosp Med* 11:494-497

Deutsche Version in:

Haseman W et al. (2007) Screening, Assessment und Diagnostik von Delirien. *Pflege* 20:191-204

Jorgensen SM, Carnahan RM, Weckmann MT (2017) Validity of the Delirium Observation Screening Scale in identifying delirium in home hospice patients. *Am J Hosp Palliative Care* 34:744-747

Koster S, Hensens AG, Oosterfeld FG et al. (2009) The Delirium Observation Screening scale recognizes Delirium early after cardiac surgery. *Eur J Cardiovasc Nurs* 8:309-314

Martínez-Arnau FM, Puchades-García A, Pérez-Ros P (2023) Accuracy of Delirium Screening Tools in Older People with Cancer — A Systematic Review. *Cancers* 15(10):2807. DOI: 10.3390/cancers15102807

Neefjes ECW, Van Der Vorst MJDL, Boddaert MSA et al. (2019) Accuracy of the Delirium Observational Screening Scale (DOS) as a Screening Tool for Delirium in Patients with Advanced Cancer. *BMC Cancer* 19:160. DOI: 10.1186/s12885-019-5351-8

Scheffer AC, van Munster BC, Schuurmans MJ, de Rooij SE (2011) Assessing severity of delirium by the Delirium Observation Screening Scale. *Int J Geriatr Psychiatry* 26:284-291

Schuurmans MJ (2001) Delirium Observatie Screening (DOS) Schaal, UMC Utrecht

Schuurmans MJ, Shortridge-Baggett LM, Duursma SA (2003) The Delirium Observation Screening Scale: a screening instrument for delirium. *Res Theory Nurs Pract Spring* 17:31-50

Teale EA, Munyombwe T, Schuurmans M et al. (2018) A prospective observational study to investigate utility of the Delirium Observation Screening Scale (DOSS) to detect delirium in care home residents. *Age Aging* 47:56-61

Park J, Jeong E, Lee J (2021) The Delirium Observation Screening Scale: A Systematic Review and Meta-Analysis of Diagnostic Test Accuracy. *Clin Nurs Res* 30(4):464-473. doi: 10.1177/1054773820961234

Tab. 2.4.22: Delirerfassung

Name des Instrumentes:	Confusion Assessment Method		
Kürzel:	CAM, CAM-long		
Mittlerer Zeitaufwand:			
Entfällt für Patient*in / für Untersucher*in (inkl. Vor- u. Nachbereitung) 9 min			
Lernaufwand für Untersucher*in:			
O leicht	x mittel	O anspruchsvoll	
Durchführung durch:			
Ärzt*in, spezifisch geschulte Pflegekraft			
Geeignet für Schweregrad:			
x leicht/Früherkennung	x mittel	x schwer	
Gütekriterien:			
Sensitivität 94-100% (Originalarbeit), 94% (gepoolte Daten aus 7 qualitativ hochwertigen Studien)			
Spezifität 90-95% (Originalarbeit), 89% (gepoolte Daten aus 7 qualitativ hochwertigen Studien)			
Interrater-Reliabilität: κ 0,8-1,0 (Originalarbeit), κ 0,7-1,0 (gepoolte Daten aus 7 qualitativ hochwertigen Studien)			
Die Güteeigenschaften hängen stark von der Qualität der Erhebung ab – Sensitivität und Spezifität können u. U. auf bis zu 40% absinken. In einer Studie zur Detektion von postoperativen Delirien fand die Gruppe um K. Rockwood signifikante Unterschiede in der Sensitivität der CAM zwischen der Anwendung des Instruments durch Ärzt*innen und Pflegepersonen (100% vs. 13%).			
Begründung der Aufnahme in die Leitlinie:			
Breit international eingeführtes Instrument zur Diagnosesicherung bei V. a. Delir, empfohlen in der S1-Leitlinie „Delir und Verwirrheitszustände inklusive Alkoholentzugsdelir“.			
Limitationen / Forschungsbedarf:			
Die Güteeigenschaften hängen stark von der Qualität der Erhebung ab.			
Literatur:			
Grover S, Kate N (2012) Assessment Scales for delirium: A review. World J Psychiatr 2:58-70			
Haseman W et al. (2007) Screening, Assessment und Diagnostik von Delirien. Pflege 20:191-204			
Hestermann U, Backenstrass M, Gekle I. et al. (2009) Validation of a German version of the Confusion Assessment Method for delirium detection in a sample of acute geriatric patients with a high prevalence of dementia. Psychopathology 42:270-276			
Inouye SK, van Dyck CH, Alessi CA et al. (1990) Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med 113:941-948			
Maschke M et al., Deutsche Gesellschaft für Neurologie (2020) Delir und Verwirrheitszustände inklusive Alkoholentzugsdelir. S1-Leitlinie AWMF-Register-Nr. 030-006			
Rolfson DB, McElhaney JE, Jhangri GS, Rockwood K (1999) Validity of the confusion assessment			

method in detecting postoperative delirium in the elderly. Int Psychogeriatr 11:431-438

Wei LA, Fearing MA, Sternberg EJ, Inouye SK (2008) The Confusion Assessment Method: a systematic review of current usage. J Am Geriatr Soc 56:823-830

2.5 Depressive Störung

Die Punktprävalenz für eine depressive Episode wird für die Altersgruppe der über 75-Jährigen mit 4,6-9,3% und für die über 85-Jährigen mit 27% angegeben [Meeks et al. 2011; Volkert et al. 2013]. Für Pflegebedürftige liegt die Punktprävalenz bei 12,4 bis 14,4% [Blazer 1987]. Im höheren Lebensalter spielen subsyndromale Depressionen eine besondere Rolle [Wolter 2016].

Die Diagnose einer depressiven Episode wird nach dem Diagnostic and Statistical Manual (DSM-V) als Gold-Standard im Rahmen eines klinischen Interviews gestellt. Dabei müssen wenigstens fünf der neun Hauptkriterien – gedrückte Stimmung, Freudlosigkeit, Insomnie oder Hypersomnolenz, Agitiertheit oder Apathie, Appetitverlust, Energielosigkeit, Gefühle von Schuld oder Wertlosigkeit, eingeschränktes Denkvermögen oder Konzentrationsschwäche, Todeswunsch - für die Diagnosestellung erfüllt sein.

Grundsätzlich können Messinstrumente zur Feststellung einer depressiven Episode immer nur Hinweise geben, nie das umfassende ärztliche Gespräch mit Anamnese und ggf. eine umfassende psychiatrische Untersuchung einschließlich einer teambasierten Beobachtung ersetzen. Immer sollte auch die Frage der Lebensmüdigkeit, des Todeswunsches und der Suizidalität angesprochen werden.

Für die Feststellung einer depressiven Episode bei geriatrischen Patient*innen sind altersadäquate Instrumente erforderlich [Jeste 2005], da alte depressive Menschen seltener suizidale Gedanken oder Dysphorie, sondern häufiger Hoffnungslosigkeit, Hilflosigkeit oder körperliche Symptome beklagen [Gallo 2005]. Bei den Instrumenten zur Erfassung einer depressiven Episode, die keine somatischen Symptome beinhalten, besteht für alte Menschen eine geringere Sensitivität [Drayer 2005], umgekehrt kann das Vorliegen rein somatisch bedingter Symptome zu falsch positiven Ergebnissen führen, wenn diese in die Bewertung einfließen.

Zur Erfassung einer depressiven Episode im höheren Lebensalter ist kein optimales Instrument verfügbar. Eine rezente Analyse [Perez 2025] kommt zu dem Ergebnis, dass viele gängige Selbstauskunftsverfahren zur Erfassung depressiver Störungen nach modernen Maßstäben eine unzureichende Qualität für wissenschaftliche Fragestellungen aufweisen. Dazu zählen auch die in dieser Leitlinie vorgestellten Instrumente PHQ-9 und HADS (s. u.).

Beobachtung, Selbstbericht, Fremdanamnese oder psychophysiologische Testung haben jeweils Vor- und Nachteile. Daher gilt die Kombination mehrerer Untersuchungsmethoden als das beste Vorgehen [Eid 2006]. Dies gilt insbesondere für Menschen mit kognitiven Problemen, daher sollte das kognitive Assessment vor dem der Stimmung erfolgen. Bei schweren Beeinträchtigungen der kognitiven und/oder sprachlichen Fähigkeiten ist die (ausschließliche) Selbstbeurteilung der Patient*innen nicht sinnvoll.

Im Folgenden werden einige bei älteren Menschen gut einsetzbare Instrumente aufgeführt.

Der **2-Fragen Test nach Whooley** (Tabelle 2.5.1) dient dem Assessment der Stufe 2a. Er erfasst zwei Hauptkriterien und wird in der S3-Leitlinie Nationale VersorgungsLeitlinie Unipolare Depression als zeitökonomisches Verfahren zum Erfassen einer Depression bei Menschen mit Risikofaktoren empfohlen. Die beiden Fragen lauten:

1. Fühlten Sie sich im letzten Monat häufig niedergeschlagen, traurig bedrückt oder hoffnungslos?
2. Hatten Sie im letzten Monat deutlich weniger Lust und Freude an Dingen, die Sie sonst gerne tun?

Whooley bewertete bereits ein einziges „Ja“ als Auffälligkeit, die o.g. Leitliniengruppe sieht nur bei Bejahen beider Fragen den Verdacht auf das Vorliegen einer Depression. In diesem Fall sowie bei anamnestischen/klinischen Hinweisen auf Depression sollte ein weiteres Instrument eingesetzt werden bzw. sollten die formalen Diagnosekriterien erfasst werden. Die Diagnose einer Depression nach ICD-10 kann nur durch explizite Erhebung aller Haupt- und Nebensymptome gestellt werden.

Die **Geriatrische Depressionsskala (GDS)** (Tabelle 2.5.2) ist eine klinische Selbstbeurteilungsskala. Die am häufigsten verwendete Form mit 15 Items (GDS 15) wurde ohne Qualitätsverlust aus einer ursprünglich 30 Item umfassenden Version entwickelt. Die einzelnen Items sind dichotomisiert mit ja oder nein zu beantworten. Je mehr Punkte erreicht werden, umso wahrscheinlicher liegt eine depressive Episode vor. Der Cut-off liegt bei 5/6 Punkten. [Yesavage 1982]. Um den Zeitaufwand für die Durchführung der GDS zu reduzieren und so die Akzeptanz für deren Anwendung zu erhöhen, wurden weitere verkürzte Versionen entwickelt. Auch Kurzversionen mit 1 bis 10 Items sind verfügbar [Pocklington et al. 2016]. In einer Untersuchung an deutschen Altenheimbewohner*innen war eine 8-Item-Version ähnlich aussagefähig wie die 15-Item-Version. Eine 4-Item-Version mit 54% Sensitivität bei einem Cut-off von 1/2 wurde als nicht empfehlenswert eingeordnet [Allgaier 2011], legt man jedoch 0/1 als Cut-off fest, so steigt die Sensitivität auf 88%.

Beim Assessment der Stimmung geriatrischer Patient*innen bietet es sich an, bei Personen ohne klinischen/anamnestischen Hinweis auf Depression die GDS-5 unter Verwendung der Items 1, 4, 8, 9 und 12 der GDS-15 einzusetzen (Stufe 2a). Bei maximal einem Punkt ist keine weitere Untersuchung notwendig, ansonsten wird z. B. die GDS-15 komplett erhoben (Stufe 2b), um den Anteil falsch positiver Ergebnisse zu reduzieren.

Der **WHO-5-Wohlbefindens-Index (WHO-5)** (Tabelle 2.5.3) ist ein kurzer Fragebogen zur Selbstbeurteilung des globalen Wohlbefindens. Wohlbefinden gilt als ein akzeptiertes Kriterium für geistige Gesundheit [Jahoda 1958]. Die Proband*in wird gebeten, bei der Beantwortung den Zeitraum der letzten 14 Tage zu bewerten. Die Items werden dabei auf einer sechsstufigen Skala von 0 bis 5 bewertet, wobei höhere Werte einer positiven Antwort entsprechen. Der Gesamtscore (0 bis 25) wird auf 100% normiert, also mit 4 multipliziert. Ein Wert von unter 50% gilt als auffällig [Jahoda 1958].

Für geriatrische Patient*innen liegen aktuell 16 Studien zur Anwendung des WHO-5 vor. Fünf davon mit insgesamt 1134 Proband*innen untersuchten die Gütekriterien des WHO-5 zur Erfassung einer depressiven Episode [Allgaier 2013, Barua 2010]. In einer Studie erfolgte der direkte Vergleich des WHO-5 mit der GDS-15. Dabei fand sich bei einer

Prävalenz einer depressiven Episode von 28,3% in der ROC-Analyse kein signifikanter Unterschied zwischen beiden Skalen. Die Sensitivität des WHO-5 betrug 92% für einen Cut-off von 13/12 und war deutlich höher als die Sensitivität der GDS 15 mit 69% bei einem Cut-off-Wert von 7/8. Der WHO-5 war insbesondere bei milden Formen einer depressiven Episode deutlich sensibler als die GDS-15 (92% vs. 54%). Die Spezifität war für den WHO-5 und die GDS-15 mit 88% vs. 79% vergleichbar [Allgaier 2013]. Bei Heimbewohner*innen fanden sich für den WHO-5 Index eine Sensitivität von 92% und eine Spezifität von 74%. Die entsprechenden Gütekriterien für die GDS-15 lauten 81% und 67% [Allgaier 2011].

Der WHO-5 ist in der Durchführbarkeit ökonomischer und wird in der AWMF S3 Leitlinie für die Detektion der unipolaren Depression (Nationale Versorgungsleitlinie, Registriernummer 005) empfohlen. Als Instrument zur Erfassung einer depressiven Episode ist der WHO-5 aufgrund seiner in mehreren Studien belegten hohen Sensitivität auch für geriatrische Patient*innen sehr geeignet [Topp 2015]. Die verfügbaren umfangreichen Ergebnisse sprechen für seinen routinemäßigen Einsatz für die Detektion einer depressiven Störung bei geriatrischen Patient*innen.

Der neun Items umfassende **Depressionsteil des Patient Health Questionnaire (PHQ-9)** (Tabelle 2.5.4) fragt nach der Häufigkeit des Auftretens jener Symptome, die gemäß Diagnostic and Statistical Manual of Mental Disorders eine Depression definieren, bezogen auf die letzten beiden Wochen. Vierstufig bewertet (von 0 für die Angabe „überhaupt nicht“ bis 3 Punkte für „beinahe jeden Tag“), errechnet sich ein Score mit maximal 27 Punkten, der dem Instrument Änderungssensitivität verleiht. Bei 5-9 Punkten wird eine Wiederholung des Assessments nach mindestens zwei Wochen empfohlen, 10-14 Punkte sprechen für eine leichte, 15-19 Punkte für eine mittelgradige, 20-27 Punkte für eine schwere Depression. Ein Vorteil gegenüber der GDS ist das direkte Ansprechen auf suizidale Gedanken. Eine Kurzform (**PHQ-2**) besteht aus den ersten beiden Fragen der PHQ-9.

Die **Depression im Alter-Skala (DIA-S)** (Tabelle 2.5.5) ist ein zeitökonomisches Instrument mit sehr guten Güteeigenschaften, das in Deutschland besonders für geriatrische Patient*innen entwickelt wurde. Die zehn Fragen sind einfacher formuliert als die Items der GDS-15 und das Auswertungsschema mit JA-NEIN-Antwortformat macht die Skala leicht handhabbar. Durch eine sorgfältige Differenzierung zwischen den Symptomen depressiver Störungen und den Begleiterscheinungen im Kontext körperlicher Erkrankungen ist die DIA-S jedoch inhaltlich besser für Patient*innen geeignet, die unter Multimorbidität leiden bzw. (teil-)stationär versorgt werden.

Die **Hospital Anxiety and Depression Scale (HADS)** (Tabelle 2.5.6) umfasst 14 Fragen und wurde als Selbstbeurteilungsbogen zur Erfassung von Angst (7 Fragen) und Depression (7 Fragen) entwickelt [Zigmond et al. 1983]. Die einzelnen Fragen werden mit 0-3 Punkten bewertet. Die HADS verzichtet bewusst auf die Erfragung somatischer Symptome. Der Summenscore umfasst für die beiden Bereiche Angst bzw. Depression jeweils 0-21 Punkte.

Die **Montgomery- Åsberg Depression Rating Scale (MADRS)** (Tabelle 2.5.7) als Fremdbeurteilungs-Instrument der Stufe 2b bietet die Möglichkeit, neben den Angaben der Patient*innen auch ihre Mimik sowie von Mitgliedern des geriatrischen Teams oder An- und Zugehörigen erhaltene Informationen in die Bewertung einfließen zu lassen.

Neben dem Einsatz bei kognitiv oder sprachlich beeinträchtigten Personen bietet die MADRS besonders dann eine wertvolle Möglichkeit, die Emotionslage zu dokumentieren, wenn sich auffällige Diskrepanzen zwischen der Sicht des geriatrischen Teams und/oder Angehörigen und der Selbstbeurteilung zeigen. Die 61-stufige Skala bildet Änderungen der Ausprägung depressiver Symptome sensitiv ab.

Tab. 2.5.1: Depressive Störung

Name des Instrumentes:	2-Fragen Test nach Whooley (Whooley Questions)		
Kürzel:	–		
Mittlerer Zeitaufwand: Für Patient*in 1 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 2 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch: Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☐ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien: Trabert (2025): n = 248, 83 J, 63% F, 2-Fragen Test und GDS-15 (s. u.) in den 24 Stunden nach stationärer Krankenhausaufnahme geriatrischer Patient*innen, MADRS (s. u.) als Goldstandard (hier ab 20 Punkten als positiv im Sinne moderat depressiver Symptome bewertet). 2-Fragen Test mit Cut-off 0/1 Sensitivität 95%, Spezifität 49%, negativer prädiktiver Wert 0,99. GDS-15 Sensitivität 57%, Spezifität 62%. Barthel-Index und MMSE beeinflussten die Resultate der GDS-15, nicht jedoch die des 2-Fragen Tests. Bosanquet (2015): Studienpopulation n = 766, ≥75 J, 46,4% F, keine kognitive Beeinträchtigung bekannt, 17 Hausarztpraxen. Bei Cut-off 0/1 Sensitivität 94,3% (95%CI 80,8–99,3%), Spezifität 62,7% (95%CI 59,0–66,2%), negativer prädiktiver Wert 0,996, Odds Ratio 27,68. Wurde keine der beiden Fragen bejaht, lag mit einer Wahrscheinlichkeit von 99,6% keine Depression vor. <u>Whooley (1997)</u> : Studienpopulation n = 536 (davon n = 97 mit Depression gemäß 20-minütigem diagnostischen Interview), 53 ±14 J, 3% F, konsekutive Patient*innen einer Bereitschaftsambulanz. Bei Cut-off 0/1 Sensitivität für das Erkennen einer Major Depression 96%, Spezifität 57%.			
Begründung der Aufnahme in die Leitlinie: Mit minimalem Aufwand erhhebbares sensitives Instrument zur Erfassung von Therapiebedarf.			
Limitationen / Forschungsbedarf: Die Gütekriterien sollten an stationär behandelten geriatrischen Patient*innen überprüft			

werden und um die Sensitivität bei der Erfassung leichterer Depressionen ergänzt werden.

Literatur:

Bosanquet K, Mitchell N, Gabe R et al. (2015) Diagnostic accuracy of the Whooley depression tool in older adults in UK primary care. J Affect Disord. 182:39-43. DOI: 10.1016/j.jad.2015.04.020.

Trabert J, Bauer E, Golbach R et al. (2025) Simplified screening for depression in acutely hospitalized geriatric patients: Comparison of the Two-Item Whooley Questions with the Geriatric Depression Scale-15. Int J Geriatr Psychiatry 40(4):e70083. DOI: 10.1002/gps.70083

Whooley MA, Avins AL, Miranda J et al. (1997) Case-finding instruments for depression. Two questions are as good as many. J Gen Intern Med 12:439-445

Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Unipolare Depression – Langfassung, Version 3.2. 2022. DOI: 10.6101/AZQ/000505

Tab. 2.5.2: Depressive Störung

Name des Instrumentes:	Geriatrische Depressionsskala		
Kürzel:	GDS		
Mittlerer Zeitaufwand:			
Für Patient*in und Untersucher*in je nach Länge der verwendeten Version			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Arzt*in, Psycholog*in, Ergotherapeut*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien:			
Park & Kwak (2021): Systematischer Review mit Meta-Analyse der Daten aus 31 Studien aus 2000-2019 mit n = 8.897. Sensitivität 80%, Spezifität 79%, AUC 0,89.			
Sheikh (1986): Die GDS-15 verfügt bei einem Cut-Off-Wert von 5/6 über eine Sensitivität für eine majore Depression von 96%, für eine minore Depression von 81% bei einer Spezifität von 63%. Die interne Konsistenz beträgt 0,79 (Cronbachs α).			
Hoyl (1999): Eine Variante der GDS-5 mit den Items 1, 4, 8, 9 und 12 der ursprünglichen GDS-15 zeigt bei einer auffälligen Antwort mit 98% die höchste Sensitivität und eine Spezifität von 88%. Die ROC Analyse ergab eine AUC von 0,94 für diese GDS-5 Variante und eine AUC von 0,91 für die GDS-15. Die Übereinstimmung zwischen dieser Variante der GDS-5 und dem Vorliegen einer Depression war hoch (κ 0,81).			
Begründung der Aufnahme in die Leitlinie:			
Von der AGAST empfohlenes, international breit eingesetztes Instrument. Einige			

Kurzformen zeigen bessere Gütekriterien als die am häufigsten verwendete 15-Item-Version.

Limitationen / Forschungsbedarf:

Die GDS ist gemäß Weeks (2003) bei MMSE unter 15, gemäß Debruyne (2009) bereits bei leichter Demenz nicht geeignet (bei MMSE 21,9 +/- 3,0 nur geringe Korrelation mit Cornell-Skala, dem Goldstandard für die Erfassung von Depression bei Demenz). Zudem sollte die GDS nicht zu früh erhoben werden, da bei so intimen Fragen zunächst eine Vertrauensbasis zwischen Patient*in und geriatrischem Team entstehen muss [Lezak 2012].

Literatur:

Allgaier AK, Kramer D, Saravo B. et al (2013) Beside the Geriatric Depression Scale: the WHO-Five Well-being Index as a valid screening tool for depression in nursing homes. Int J Geriatr Psychiatry 28:1197-1204

Allgaier AK, Kramer D, Mergl R. et al. (2011) Validität der Geriatrischen Depressionsskala bei Altenheimbewohnern: Vergleich von GDS-15, GDS-8 und GDS-4. Psychiatr Prax. 38:280-286

Debruyne H, Van Buggenhout M, Le Bastard N et al. (2009) Is the geriatric depression scale a reliable screening tool for depressive symptoms in elderly patients with cognitive impairment? Int J Geriatr Psychiatry 24:556-562

Gauggel S, Birkner B (1999) Validity and reliability of a German version of the Geriatric Depression Scale (GDS). Zeitschrift für Klinische Psychologie-Forschung und Praxis 28:18–27

Hoyl MT, Alessi CA, Harker JO et al. (1999) Development and testing of five-item version of the Geriatric Depression Scale. J Am Geriatr Soc 7:873-878

Park SH, Kwak MJ (2021) Performance of the Geriatric Depression Scale-15 with older adults aged over 65 years: An updated review 2000-2019. Clinical Gerontologist 44(2):83–96. DOI: 10.1080/07317115.2020.1839992

Trabert J, Bauer E, Golbach R et al. (2025) Simplified screening for depression in acutely hospitalized geriatric patients: Comparison of the Two-Item Whooley Questions with the Geriatric Depression Scale-15. Int J Geriatr Psychiatry 40(4):e70083. DOI: 10.1002/gps.70083

Weeks SK, McGann PE, Michaels TK, Penninx B (2003) Comparing various short-form Geriatric Depression Scales leads to the GDS 5/15. J Nurs Schol 35:133-137

Yesavage JA, Brink TL, Rose TL et al. (1982) Development and validation of a geriatric depression screening scale. A preliminary report. J Psychiatr Res 17:37-49

Tab. 2.5.3: Depressive Störung/ Wohlbefinden

Name des Instrumentes:	WHO-5-Wohlbefindens-Index	
Kürzel:	WHO-5	
Mittlerer Zeitaufwand:		
Für Patient*in 5 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 7 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Ärzt*in, Psycholog*in, Ergotherapeut*in, spezifisch geschultes Fachpersonal		

Geeignet für Schweregrad:		
x leicht/Früherkennung	x mittel	x schwer
Gütekriterien:		
Kliem (2025): n = 2515, zufällig ausgewählt in Deutschland, davon ab 75 J 138 F, 99 M. Hohe interne Konsistenz (Cronbachs α 0,95) und Korrelation mit Instrumenten, die Depression, Angst, somatische Symptome und Einsamkeit abbilden. Angabe von Perzentilen für verschiedene Altersgruppen (ab 75 J 16. Perzentile 10,5 von 100 Punkten (F 11,6 P, M 9,1 P). Allgaier (2013) Der WHO-5 Index war insbesondere hinsichtlich der Detektion von milden Formen einer depressiven Episode der GDS-15 deutlich überlegen.		
Begründung der Aufnahme in die Leitlinie:		
Ermöglicht die Detektion einer depressiven Störung indirekt durch Mangel an positiv belegten Stimmungs-Episoden, geeignet auch bei Abwehr des Gedankens, evtl. depressiv zu sein.		
Limitationen / Forschungsbedarf:		
Kognitiv gestörte Patient*innen haben oft Schwierigkeiten, sich den Beurteilungszeitraum von 2 Wochen zu vergegenwärtigen.		
Literatur:		
Deutsche Version unter https://www.psychiatri-regionh.dk/who-5/Documents/WHO5_German.pdf Allgaier AK, Kramer D, Saravo B et al. (2013) Beside the Geriatric Depression Scale: The WHO-Five Well-being Index as a valid screening tool for depression in nursing homes. Int J Geriatr Psychiatry 28:1197–1204 DOI: 10.1002/gps.3944 Barua A, Kar N (2010) Screening for depression in elderly Indian population. Ind J Psychiatr 52:150-153 DOI: 10.4103/0019-5545.64595 Kliem S, Lohmann A, Fischer S et al. (2025) Psychometric evaluation and updated community norms of the WHO-5 well-being index, based on a representative German sample. Front Psychol 16:1592614. DOI: 10.3389/fpsyg.2025.1592614Topp CW, Østergaard SD, Søndergaard S, Bech P (2015) The WHO-5 Well-Being Index: A Systematic Review of the Literature. Psychother Psychosom 84:167-176 WHO (1998) Wellbeing Measures in Primary Health Care/The Depcare Project. WHO Regional Office for Europe: Copenhagen		

Tab. 2.5.4: Depressive Störung

Name des Instrumentes:	Patient Health Questionnaire, 9-Item Depressions-Modul
Kürzel:	PHQ-9
Mittlerer Zeitaufwand: Für Patient*in 6min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 8 min	
Lernaufwand für Untersucher*in:	

x leicht	O mittel	O anspruchsvoll
Durchführung durch: Ärzt*in, Psycholog*in		
Geeignet für Schweregrad:		
x leicht/Früherkennung	x mittel	x schwer
Gütekriterien: Tack (2023): 50 stationäre onkologische Patient*innen im Alter >70 J (Median 76 J), 32% F, 68% mit professioneller ambulanter Pflege. Goldstandard SCID-5 strukturiertes Interview. Die höchste diagnostische Güte gemäß AUC der ROC erreichte die PHQ-2 (0,90; bei Cut-off 1/2 Sensitivität 88,9%, Spezifität 87,8%, positiver prädiktiver Wert 61,5%, negativer 97,3%) vor der PHQ-9 (0,87) vor der GDS-4 (0,72; Items 1, 3, 6 und 7 der GDS-15) vor der GDS-15 (0,63). Sensitivität und Spezifität für die Erkennung einer Major Depression bei einem Cut-off von 9/10 jeweils 0,88 [Kroenke 2001]. Interne Konsistenz: Cronbachs α 0,88 [Gräfe 2004]. Für die telefonische Durchführung validiert [Fine 2013]. Für die Kurzform PHQ-2 werden bei einem Cut-off von 2/3 eine Sensitivität und Spezifität von 86% bzw. 78% angegeben und bei einem Cut-off von 3/4 61% bzw. 91% [Arroll et al. 2010]. Begründung der Aufnahme in die Leitlinie: Alternative zur GDS, für den ambulanten Bereich als änderungssensitives Instrument empfohlen in der Nationalen Versorgungsleitlinie Unipolare Depression. Limitationen / Forschungsbedarf: Da Schwierigkeiten mit der zeitlichen Orientierung und damit auch mit der Angabe von Häufigkeiten oft bereits früh im Verlauf einer demenziellen Entwicklung auftreten, ist die Zuverlässigkeit in dieser Situation eingeschränkt. Literatur: Arroll B, Goodyear-Smith F, Crenle S, Gunn J, Kerse N, Fishman K, Hatcher S (2010) Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. Ann Fam Med 8:348-353 DOI: 10.1370/afm.1139 Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Unipolare Depression – Langfassung, Version 3.2. 2022. DOI: 10.6101/AZQ/000505 Fine TH, Contractor AA, Tamburrino M et al. (2013) Validation of the telephone administered PHQ-9 against the in-person administered SCID-I major depression module. J Affect Disord 150:1001–1007 Gräfe K, Zipfel S, Herzog W, Löwe B (2004) Screening psychischer Störungen mit dem „Gesundheitsfragebogen für Patienten (PHQ-D)“ Ergebnisse der deutschen Validierungsstudie. Diagnostica 50:71–181 DOI: 10.1026/0012-1924.50.4.171 Kroenke K, Spitzer RL, Williams JB (2001) The PHQ-9. Validity of a brief depression severity measure. J Gen Intern Med 16:606–613 Kroenke K, Spitzer RL (2002) The PHQ-9: A new depression diagnostic and severity measure. Psychiatric annuals 32:1-7 Perez NB, Sierra PG, Taylor B, Fletcher J (2025) Selecting a depression measure for research: A		

critical examination of five common self-report scales. J Affect Disord 26:120542. DOI: 10.1016/j.jad.2025.120542

Spangenberg L, Brähler E, Glaesner H (2012) Wie gut eignen sich verschiedene Versionen des Depressionsmoduls des Patient Health Questionnaire zur Identifikation depressiver Personen in der Allgemeinbevölkerung? Z Psychosom Med Psychother 58:3-10

Spitzer RL, Kroenke K, Williams JB et al. (1999) Validation and utility of a self-report version of PRIME-MD. The PHQ Primary Care Study. JAMA 282:1737–1744

Tack L, Maenhoudt AS, Ketelaars L et al. (2023) Diagnostic Performance of Screening Tools for Depressive Symptoms in Vulnerable Older Patients with Cancer Undergoing Comprehensive Geriatric Assessment (CGA): Results from the SCREEN Pilot Study. Curr Oncol 30:1805–1817. <https://doi.org/10.3390/curroncol30020140>

Tab. 2.5.5: Depressive Störung

Name des Instrumentes:	Depression im Alter-Skala		
Kürzel:	DIA-S		
Mittlerer Zeitaufwand:			
Für Patient*in 8 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 10 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Ärzt*in, Psycholog*in, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien:			
Wunner (2022): Studienpopulation n = 375 Patient*innen einer geriatrischen Tagesklinik in Deutschland, 81,1 ± 5,4 J, 74% F, 39% mit Demenzdiagnose. Erhebung durch geschultes Pflegefachpersonal bei Aufnahme, Goldstandard psychologisches halbstandardisiertes Interview zu den Diagnosekriterien gemäß DSM 5. Im Vergleich zur GDS-15 Sensitivität höher (bei Cut-off 3/4 92%, bei Cut-off 4/5 81%, GDS-15 bei Cut-off 5/6 78%), jedoch Spezifität niedriger (57% bzw. 70% gegenüber 77% bei der GDS-15). Entgegen Erwartung bei Demenzdiagnose und Cut-Off 4/5 niedrigere Sensitivität und Spezifität als GDS-15.			
Jung (2019): Studienpopulation n = 385 ohne schwere Demenz (Mini Mental State Examination >14), davon n = 45 mit majorer, 129 mit minorer depressiver Störung, 70,2 ± 6,0 J, 60,0% F. Ambulanzen von 8 Universitäten. Korrelation der koreanischen Übersetzung der DIA-S aus dem Englischen mit der GDS r = 0,853, dem PHQ 0,739 und der MADRS 0,772 festgestellt, Cut-off von 3/4 hinsichtlich jeder depressiven Störung bestätigt. Die Autoren empfehlen für die Erfassung einer majoren depressiven Episode einen Cut-off von 4/5 (Sensitivität 91%, Spezifität 75%).			
<u>Heidenblut (2010)</u> : Studienpopulation n = 331 ohne schwere Demenz (Mini Mental State			

Examination >15), 151 davon mit Hinweis auf Depressivität (Montgomery Asberg Depression Rating Scale >12), 80,7 ±7,6 J, 73% F, 3 geriatrische Klinik-Abteilungen. Bester Cut-off bei 3/4. Sensitivität (82%), Spezifität (79%), Korrelation mit der MADRS (0,728), externe und interne Trennschärfe der Items sowie interne Konsistenz (0,84) jeweils höher als die der GDS-15.

Begründung der Aufnahme in die Leitlinie:

Sprachlich deutlich einfacher als die GDS gestaltetes Instrument.

Limitationen / Forschungsbedarf:

Bislang liegt im deutschsprachigen Raum nur eine Validierungsstudie im direkten Vergleich mit der GDS-15 bei stationären geriatrischen Patient*innen vor, wobei die MADRS hierbei als Goldstandard verwendet wurde. Eine weitere Erprobung der DIA-S im teilstationären und ambulanten Setting ist wünschenswert.

Literatur:

Deutsche Version unter <https://www.hf.uni-koeln.de/data/gerontologie/File/DIASkalaneu.pdf> (Zugriff 10.07.2022)

Heidenblut S, Zank S (2014) Screening for depression with the Depression in Old Age Scale (DIA-S) and the Geriatric Depression Scale (GDS15). Diagnostic accuracy in a geriatric inpatient setting. Gero Psych 27:41–49

Heidenblut S, Zank S (2010) Entwicklung eines neuen Depressionsscreenings für den Einsatz in der Geriatrie. Die „Depression-im-Alter-Skala“ (DIA-S). Z Gerontol Geriatr. 43:170-176 DOI 10.1007/s00391-009-0067-z

Jung YE, Kim MD, Bahk WM, Woo YS, Nam B, Seo JS, Jang SH, Sung HM, Shim IH, Yoon BH, Kim JS, Kwon YJ. (2019) Validation of the Korean Version of the Depression in Old Age Scale and Comparison with Other Depression Screening Questionnaires Used in Elderly Patients in Medical Settings. Clin Psychopharmacol Neurosci. 17(3):369-376

Wunner S, Stemmler M, Masuch J et al. (2022) Screening for depression in old age. comparison of the geriatric depression scale and the depression in old age scale. Z Gerontol Geriatr 55(1):44-50. DOI: 10.1007/s00391-021-01949-w

Tab. 2.5.6: Depressive Störung/Angststörung

Name des Instrumentes:	Hospital Anxiety and Depression Scale		
Kürzel:	HADS		
Mittlerer Zeitaufwand:			
Für Patient*in 8 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 12 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch:			
Ärzt*in, Psycholog*in			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☒ schwer	

Gütekriterien:

Auch für ältere Menschen (65-80 Jahre) validiert, gute interne Konsistenz (Cronbachs α 0,87 für Depression, 0,81 für Angst [Djukanovic et al. 2017]. Folgende Bewertung wird angegeben: 0-7 Punkte unauffällig, 8-10 Punkte grenzwertig, mehr als 10 Punkte auffällig [Zigmond et al. 1983]. Sensitivität und Spezifität jeweils 80%.

Wu (2021): Meta-Analyse, kumuliert n = 22 574, meist halb-strukturiertes Interview, Alter überwiegend unter 60 J, empfehlen Cut-off 6/7, Sensitivität dann 82%, Spezifität 78%.

Begründung der Aufnahme in die Leitlinie:

Getrennte Auswertung von Symptomen einer Angststörung möglich, gehört zu den in der Nationalen Versorgungsleitlinie Unipolare Depression und S3-Leitlinie Diagnostik und Therapie Bipolarer Störungen empfohlenen Instrumenten.

Limitationen / Forschungsbedarf:

Geeignet, aber bisher wenig in der Geriatrie eingesetzt.

Literatur:

Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Unipolare Depression – Langfassung, Version 3.2. 2022. DOI: 10.6101/AZQ/000505

DGBS e.V. und DGPPN e.V. (2019) S3-Leitlinie Diagnostik und Therapie Bipolarer Störungen. Langversion. AWMF-Register-Nr. 038-019

Djukanovic I, Carlsson J, Årestedt K (2017) Is the Hospital Anxiety and Depression Scale (HADS) a valid measure in a general population 65–80 years old? A psychometric evaluation study. Health and Quality of Life Outcomes 15:193. DOI 10.1186/s12955-017-0759-9

Herrmann Ch, Buss U (1994) Vorstellung und Validierung einer deutschen Version der "Hospital Anxiety and Depression Scale" (HAD-Skala); ein Fragebogen zur Erfassung des psychischen Befindens bei Patienten mit körperlichen Beschwerden. Diagnostica 40:143-154

Perez NB, Sierra PG, Taylor B, Fletcher J (2025) Selecting a depression measure for research: A critical examination of five common self-report scales. J Affect Disord 26:120542. DOI: 10.1016/j.jad.2025.120542Wu Y, Levis B, Sun Y et al. (2021) Accuracy of the Hospital Anxiety and Depression Scale Depression subscale (HADS-D) to screen for major depression: systematic review and individual participant data meta-analysis. DOI: <https://doi.org/10.1136/bmj.n972>

Zigmond AS, Snaith RP (1983) The hospital anxiety and depression scale. Acta Psychiatr Scand. 67:361–370 DOI: 10.1111/j.1600-0447.1983.tb09716.x

Tab. 2.5.7: Depressive Störung

Name des Instrumentes:	Montgomery-Åsberg Depression Rating Scale	
Kürzel:	MADRS	
Mittlerer Zeitaufwand:		
Für Patient*in 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 15 min		
Lernaufwand für Untersucher*in:		
☐ leicht	☒ mittel	☐ anspruchsvoll

Durchführung durch:

Ärzt*in, Psycholog*in

Geeignet für Schweregrad:

x leicht/Früherkennung	x mittel	x schwer
------------------------	----------	----------

Gütekriterien:

Atchison (2024): Systematischer Review mit Meta-Analyse von 23 Studien zur diagnostischen Genauigkeit bei gleichzeitigem Vorliegen einer Demenz. Bei Wahl des optimalen Cut-offs Sensitivität / Spezifität für die MADRS 77%/68% (für die GDS-15 65%/72%).

Seemüller (2023): n = 919, stationäre Patient*innen mit majorer Depression in 12 deutschen psychiatrischen Kliniken, kein Einschluss bei Alter >65 J (!), mittleres Alter 46 ± 12 J, 62% F. Interne Konsistenz hoch mit Cronbachs α 0,92. Die Items „berichtete Traurigkeit“, „sichtbare Traurigkeit“ und „Gefühllosigkeit“ korrelierten am engsten mit dem Gesamtscore, vier Faktoren identifiziert.

Smalbrugge (2008): Studienpopulation (Daten aus der Amsterdam Groningen Elderly Depression Studie) n = 350, davon n = 27 Personen mit majorer, n = 47 Personen mit minorer Depression gemäß DSM-IV, 79,3 ± 8,3 (55 – 99) J, 68,9% F, Bewohnende in 14 Pflegeheimen (Niederlande) ohne schwere Demenz (Mini Mental State Examination >14). Über die Effektstärke abgebildete Änderungssensitivität für keine/leichte Depression (minor depression) (1,94) und minor/major depression (1,61) fast doppelt so hoch wie bei der GDS-15.

Schmidtke (1988): Interne Konsistenz hoch mit Cronbachs α 0,86.

Montgomery (1979): Interrater-Reliabilität bei verschiedenen Konstellationen zweier Untersucher 0,89 – 0,97 bei Fallzahlen von n = 11 bis n = 30

Begründung der Aufnahme in die Leitlinie:

Auch bei fehlender Selbstbeurteilungsfähigkeit einsetzbar, Einbezug anamnestischer Angaben Angehöriger erlaubt, änderungssensitiver als GDS, in der Nationalen Versorgungsleitlinie Unipolare Depression und S3-Leitlinie Diagnostik und Therapie Bipolarer Störungen empfohlen.

Limitationen / Forschungsbedarf:

Geeignet, aber bisher wenig in der Geriatrie eingesetzt.

Literatur:

Atchison K, Nazir A, Wu P (2024) Depression detection in dementia: A diagnostic accuracy systematic review and meta analysis update. Health Sci Rep 7(11):e70058. DOI: 10.1002/hsr2.70058

Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Unipolare Depression – Langfassung, Version 3.2. 2022. DOI: 10.6101/AZQ/000505

DGBS e.V. und DGPPN e.V. (2019) S3-Leitlinie Diagnostik und Therapie Bipolarer Störungen. Langversion. AWMF-Register-Nr. 038-019

Montgomery SA, Åsberg M (1979) A new depression scale designed to be sensitive to change. Br J Psychiatry 134:382-389

Schmidtke A, Fleckenstein P, Moises W, Beckmann H (1988) Untersuchungen zur Reliabilität und

Validität einer deutschen Version der Montgomery-Åsberg Depression Rating Scale. Schweiz Arch Neurol Psychiat 139:51-65

Seemüller F, Schennach R, Musil R et al. (2023) A factor analytic comparison of three commonly used depression scales (HAMD, MADRS, BDI) in a large sample of depressed inpatients. BMC Psychiatry 23:548. <https://doi.org/10.1186/s12888-023-05038-7>

Smalbrugge M, Jongenelis L, Pota AM et al. (2008) Screening for depression and assessing change in severity of depression. Is the Geriatric Depression Scale (30-, 15- and 8-item versions) useful for both purposes in nursing home patients? Aging & Mental Health 12:244–248

Beispiele:

1. Patient*in anamnestisch und klinisch ohne Hinweis auf depressive Störung, kognitiv allenfalls mäßig beeinträchtigt.
→ **z. B. 2-Fragen-Test nach Whooley 0 Punkte (Stufe 2a)**
→ **derzeit keine Indikation zu weiterem Assessment**
2. Patient*in anamnestisch oder/und klinisch mit Hinweis auf mögliche depressive Störung, kognitiv allenfalls mäßig beeinträchtigt.
→ **z. B. DIA-S / GDS-15, WHO-5**
3. Patient*in kognitiv stark beeinträchtigt.
→ **Versuch des Einsatzes einer Fremdbeurteilungsskala z. B. MADRS**

2.6 Soziale Situation

Die Möglichkeiten, sein Leben so zu führen, wie man es möchte, werden nicht nur von Faktoren innerhalb einer Person bestimmt, sondern sie sind in bedeutendem Umfang auch von äußeren Bedingungen abhängig wie der sozialen Situation. Pflichtelemente des Sozialassessments im Rahmen der frührehabilitativen Komplexbehandlung [Bundesinstitut für Arzneimittel und Medizinprodukte 2026] sind die Erhebung des sozialen Umfeldes, der Wohnsituation, der häuslichen und außerhäuslichen Aktivitäten (bei stationärer Therapie: vor dem Akutereignis), vorhandener/benötigter Hilfsmittel, vorhandener/benötigter pflegerischer Unterstützung und rechtlicher Verfügungen. Weder der Sozialfragebogen **SOS Hochzirl** noch der für die Stufe 2 zu umfangreiche **Sozialfragebogen nach Nikolaus** [Nikolaus 1994] beinhalten alle notwendigen Aspekte, zurzeit werden überwiegend hausintern standardisierte Instrumente eingesetzt.

Soziale Isolation ist ein etablierter Risikofaktor für kognitiven Abbau [Kelly 2017; Kuiper 2016], Frailty [Hanlon 2024], Heimübersiedlung [Yong 2020], Mortalität [Rasulo 2005], Depression [Schwarzbach 2014], Aufnahme in Notaufnahme [Steinmiller 2015] sowie weitere Einbußen der Gesundheit und Selbsthilfefähigkeit [WHO 2021; Courtin & Knapp 2017] und krankheits- und altersrelevante Biomarker [Cudjoe 2022]. Soziale Isolation und Einsamkeit als Trigger für multiple geriatritypische Störungen werden bereits als neue „geriatric giants“ [Freedman 2020] bezeichnet.

Die **Lubben Social Network Scale (LSNS)** wird seit den 1980-er Jahren in vielen etablierten Studien als Interview oder zum Selbst-Ausfüllen bei älteren Menschen

eingesetzt und gehört heute zu den anerkanntesten Instrumenten für eine zeitökonomische und gleichzeitig verlässliche Einschätzung von sozialer Einbindung bzw. Isolation. Die LSNS wurde gemäß einem rezenten Reviewpapier [Hanlon 2024] zum Zusammenhang von sozialer Isolation und Frailty z. B. in 12 der einbezogenen Studien eingesetzt, in Deutschland in den Studien zu ActiFE [Denkinger 2010] sowie der LIFE-Adult Studie [Weitzel 2023]. Die **6-Item-Version der Lubben Social Network Scale (LSNS-6)** (Tabelle 2.7.1) enthält je drei Fragen zu Kontakten mit Familienangehörigen einerseits, Freund*innen und Nachbar*innen andererseits, die zu einer sechsstufigen Bewertung auf Item-Ebene führen. Daraus resultiert ein Summenscore von 0 bis 30 Punkten; unter 12 Punkte weisen auf soziale Isolation hin.

Tab. 2.7.1: Soziale Situation

Name des Instrumentes:	Lubben Social Network Scale - 6-Item-Version		
Kürzel:	LSNS-6		
Mittlerer Zeitaufwand:			
Für Patient*in 3 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 3 min			
Lernaufwand für Untersucher*in:			
☒ leicht		☐ mittel	☐ anspruchsvoll
Durchführung durch:			
Arzt/Ärztin, Pflegekraft, Sozialdienst, spezifisch geschultes Fachpersonal			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung		☒ mittel	☐ schwer
Gütekriterien:			
Chan (2023): n = 324, 71,7±8,6 (58-95) J, Hongkong. Interne Konsistenz Cronbachs α 0,75, 2-Faktoren-Modell.			
Myagyarmav (2019): n = 650, 66,5±6,9 J, berentete Bewohnende der Mongolei. Vergleichende Validierung der 18- und der 6-Item-Version (u. a. Einsatz von Instrumenten zur Erfassung von Depressivität, Lebensqualität, sozialer Isolation). Die LSNS-6 wurde als besser geeignet bewertet.			
Chang (2018): n = 2819, 70,8±7,7 (60-96) J, 51% F, randomisiert ausgewählte Bewohnende des Raums Peking. Interne Konsistenz Cronbachs α 0,83. Verminderte Suizidalität bei guter familiärer Vernetzung.			
Lubben (2006): Daten aus der PRO-AGE-Studie, an der ambulante ab 65-Jährige (im Mittel 74 J) ohne Unterstützungsbedarf in den Basis-Aktivitäten des täglichen Lebens in Hamburg, London und Solothurn (Schweiz) teilnahmen. Interne Konsistenz Cronbachs α = 0,83			
Goetz (2001): 100 ambulante Personen in Ulm, Hamburg oder Bern, im Mittel 83 J, 81% F. Retest-Reliabilität Cohens kappa 0,64			

Begründung der Aufnahme in die Leitlinie:

Zeitökonomisch einsetzbare Selbstbeurteilungsskala zur Erfassung der familiären und außerfamiliären sozialen Einbettung als wichtigem Faktor der Lebensqualität und Prognose.

Limitationen / Forschungsbedarf:

Nicht bei mittlerer oder schwerer Demenz geeignet, es ist jedoch die Befragung Angehöriger möglich. Die Überprüfung der Gütekriterien bei geriatrischen Patient*innen ist noch unvollständig.

Literatur:

Chan SCY, Wong CC, Huang QL, Fung CK (2023) The psychometric properties of the Lubben Social Network Scale (LSNS-6) and its associations with well-being indicators in Hong Kong older adults. *Australasian Journal on Ageing* 42(4):683-689

Chang Q, Sha F, Chan CH, Yip PSF (2018) Validation of an abbreviated version of the Lubben Social Network Scale ("LSNS-6") and its associations with suicidality among older adults in China. *PLoS One* 13(8):e0201612. DOI: 10.1371/journal.pone.0201612

Goetz SM, Stuck AE, Hirschi A et al. (2001). Test-Retest-Reliabilität eines deutschsprachigen multidimensionalen Assessmentinstruments bei älteren Personen. *Z Geront Geriatr* 34:196–206. DOI: 10.1007/s003910170064

Hanlon P, Wightman H, Politis M et al. (2024) The relationship between frailty and social vulnerability: a systematic review. *The Lancet Healthy Longevity* 5(3):e214-e226. [https://doi.org/10.1016/S2666-7568\(23\)00263-5](https://doi.org/10.1016/S2666-7568(23)00263-5)

Lubben JE (1988) Assessing social networks among elderly populations. *Family & Community Health* 11(3):42-52. <https://doi.org/10.1097/00003727-198811000-00008>

Lubben JE (1984) Lubben Social Network Scale (LSNS). APA PsycTests. <https://doi.org/10.1037/t10706-000>

Lubben J, Blozik E, Gillmann G et al. (2006). Performance of an Abbreviated Version of the Lubben Social Network Scale Among Three European Community-Dwelling Older Adult Populations. *The Gerontologist* 46:503–513. DOI: 10.1093/geront/46.4.503

Lubben JE & Gironde MW (2003a) Centrality of social ties to the health and well-being of older adults. In: Berkman B & Harooytan LK (Hrsg.) *Social work and health care in an aging world* (S. 319–350) New York: Springer

Myagmarjav S, Burnette D, Goeddeke F (2019) Comparison of the 18-item and 6-item Lubben Social Network Scales with community-dwelling older adults in Mongolia. *PLoS One* 14(4):e0215523. DOI: 10.1371/journal.pone.0215523

Weitzel EC, Glaesmer H, Hinz A. et al. (2023) Soziodemografische und soziale Korrelate selbstberichteter Resilienz im Alter – Ergebnisse der populationsbasierten LIFE-Adult-Studie. *Bundesgesundheitsbl* 66:402–409. DOI: 10.1007/s00103-023-03675-7

2.7 Schmerz

Schmerz ist ein häufiges Symptom im Alter. Etwa 70% der älteren pflegebedürftigen Menschen leiden unter chronischen Schmerzen [Dräger 2021]. Weil viele geriatrische Patient*innen allerdings nicht über ihre Schmerzen klagen („underreporting of pain“), sollten sie explizit danach gefragt werden [S3-Leitlinie Schmerzmanagement bei GERiatrischen PATieNt:innen AWMF Reg.Nr. 145-005]. Im Schmerzassessment gilt die Regel „Selbsteinschätzung vor Fremdeinschätzung“, das heißt, wann immer möglich,

sollten Betroffene zunächst selbst nach Schmerzen befragt werden („Haben Sie momentan Schmerzen?/Tut Ihnen im Moment etwas weh? Hatten Sie in den vergangenen zwei Wochen Schmerzen?“)

Sofern ausnahmsweise unter Berücksichtigung der Patientensicht trotz positiver Schmerzanamnese kein weiterer Klärungs- und/oder Therapiebedarf besteht, ist der Grund zu dokumentieren. Ansonsten soll ein Assessment der Stufe 2b erfolgen. Dies umfasst die Erhebung

1. der Schmerzintensität,
2. der Lokalisation der Schmerzen (ggf. Einzeichnen auf einem Körperdiagramm),
3. des Schmerzcharakters,
4. der Schmerzdauer (seit wann, wie oft, wie lange),
5. verstärkender oder lindernder Faktoren (inkl. der Medikation).

Der **painDETECT** Fragebogen [Freynhagen 2006] eignet sich zur Erkennung von neuropathischen Schmerzen.

Visuelle Analogskalen (VAS) zur Einschätzung der Schmerzstärke sind bei geriatrischen Patient*innen häufig nicht gut einsetzbar. Bei differenzierten Patient*innen kann eine **Numerische Rating Skala (NRS)** angewendet werden. Dabei ist ein kritisches Nachfragen zum Verständnis hilfreich (z. B. soll die Person zunächst beispielhaft angeben, welche Einstufung sie wählen würde, wenn sie keine, leichte oder schwere Schmerzen hätte). Bei eingeschränkter kognitiver Leistungsfähigkeit ist einer **verbale(n) Ratingskala (VRS)** der Vorzug zu geben.

Nur wenn das Erheben einer Selbstausskunft zu Schmerzen nicht möglich erscheint (z. B. wegen einer fortgeschrittenen Demenz), sollte auf Fremdeinschätzungs-Instrumente zurückgegriffen werden. Es ist nicht möglich, eine klare Grenze zu ziehen, bis zu welchem Grad des Absinkens der kognitiven Leistungsfähigkeit Selbstausskünfte ausreichend valide sind. In Zweifelsfällen und insbesondere bei Inkongruenz zwischen Angaben und nonverbalen Indizien soll eine zusätzliche Fremdeinschätzung unter Einbezug von Vorinformationen, Fremdanamnese und Verhalten (z. B. Schonhaltung oder Grimassieren, Lautäußerungen usw. bei Mobilisierung) erfolgen [Sirsch 2020].

Die im deutschsprachigen Raum gebräuchlichsten Fremdeinschätzungs-Instrumente sind die **Skala zur Beurteilung von Schmerzen bei Demenz (BESD-Skala)** (Tabelle 2.7.1) und das **Beobachtungsinstrument für das Schmerzassessment bei alten Menschen mit Demenz (BISAD-Skala)** (Tabelle 2.7.2).

Die **BESD-Skala** lässt sich auch ohne Vorinformationen gut in einem klinischen Setting anwenden. Insgesamt werden 5 Items (Atmung, negative Lautäußerung, Gesichtsausdruck, Körpersprache und Trost) beobachtet und mit 0-2 Punkten bewertet. Maximal können 10 Punkte erreicht werden. Die Beobachtung erfolgt jeweils in Ruhe und bei Bewegung, etwa einem Transfer.

Bei der **BISAD-Skala** handelt es sich um die deutsche Adaptation der französischen ECPA (Échelle Comportementale de la douleur pour Personnes Âgées non

communicantes) [Morello 2007]. Die Beurteilung wird unterteilt in „vor der Mobilisation“ (= Bewegung induzierender Handlung, z. B. Lagewechsel im Bett) und „während der Mobilisation“ (jeweils Intensitätsabstufung von 0 bis 4). Vor der Mobilisation werden die folgenden 4 Bereiche beurteilt: Gesichtsausdruck, spontane Ruhehaltung, Bewegung oder Mobilität und „Über die Beziehung zu anderen“. Während der Mobilisation werden die folgenden 4 Bereiche beurteilt: Ängstliche Erwartung bei der Pflege, Reaktionen während der Mobilisation, Reaktionen während der Pflege der schmerzenden Bereiche und während der Pflege vorgebrachte Klagen. Der Summenscore liegt zwischen 0 und 32.

Bei beiden o.g. Instrumenten fließt ein gewisses Maß an subjektiver Beurteilung in die Bewertung ein. Das Vorliegen von Komorbiditäten, die die Motorik beeinflussen können (z. B. Parkinsonsyndrom), kann das Assessment deutlich erschweren. Der Summenscore ist weniger als interindividuell objektives Maß zu werten, wichtig ist vielmehr der individuelle Verlauf.

Sowohl bei kognitiv unauffälligen als auch beeinträchtigten Personen ist in der Stufe 3 die Rolle psychischer Faktoren im Schmerzempfinden zu berücksichtigen und je nach vermuteter Genese der Beschwerden spezifische Diagnostik durchzuführen (z. B. ausführliche körperliche Untersuchung, Röntgendiagnostik, Messung der Nervenleitgeschwindigkeit).

Die Intervalle der Schmerzerfassung sollen individuell gewählt werden [S3-Leitlinie Schmerzmanagement bei Geriatrischen Patient:innen in allen Versorgungssettings (GeriPAIN)]. Einflussfaktoren auf die Häufigkeit der Messungen sind u. a. die Schwere der Schmerzen, die Bedürfnisse der Patient*in, ob eine stabile oder eine instabile Schmerzsituation vorliegt, aber auch die Art der Einrichtung, in der Schmerz gemessen wird. In einem Krankenhaus sollte ein Schmerzassessment (vorwiegend die Dokumentation der Schmerzintensität) mindestens zweimal täglich erfolgen (etwa einmal pro Pflegeschicht), bis eine ausreichende Schmerzreduktion erreicht wurde. Dies ist der Fall, wenn der Patient sagt, dass seine Schmerzen ausreichend gelindert sind und seine körperliche, psychische und soziale Funktionsfähigkeit wiederhergestellt ist. Feste Cut-off-Werte (z. B. NRS < 4), werden heute eher kritisch betrachtet. Entscheidend ist vor allem die intraindividuelle Schmerzreduktion. Hierbei sollten die individuellen Therapieziele Berücksichtigung finden. Auch jedes neue einschneidende Ereignis – wie z. B. ein Sturz – sollte zu einer Schmerzmessung führen.

Tab. 2.7.1: Schmerz

Name des Instrumentes: Kürzel:	Beurteilung von Schmerzen bei Demenz BESD	
Mittlerer Zeitaufwand: Entfällt für Patient*in / für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min		
Lernaufwand für Untersucher*in:		
☐ leicht	☒ mittel	☐ anspruchsvoll

Durchführung durch:

Ärzt*in, Pflegekraft, Ergotherapeut*in, Physiotherapeut*in

Geeignet für Schweregrad:

☐ leicht/Früherkennung	☒ mittel	☐ schwer
--	---	------------------------------

Gütekriterien:

Die deutsche Version wurde an 99 Pflegeheimpatient*innen mit demenzieller Entwicklung evaluiert. Gute interne Konsistenz (Cronbach`s alpha = 0,85), Interrater-Reliabilität 0,80, Retest-Reliabilität 0,90.

BESD-I-Studie: n = 12 leicht bis mittelschwer kognitiv eingeschränkte Patienten (im Mittel 84,3 J), deutlicher Einfluss der Schmerzmedikation auf das in der BESD beurteilte Verhalten als Hinweis auf die Konstruktvalidität.

BESD-II-Studie: n = 44 stark kognitiv eingeschränkte Patient*innen, doppelt-verblindete, randomisierte, Placebo-kontrollierte Studie. Kein signifikanter Unterschied in der Reduktion laut BESD Schmerz zugeordneten Verhaltens zwischen Patienten, die mit einem Verum bzw. einem Placebo behandelt wurden.

Begründung der Aufnahme in die Leitlinie:

Zur Zeit der Aufnahme das am breitesten angewandte Instrument zur Beurteilung von Schmerzen bei Patient*innen mit Demenz.

Limitationen / Forschungsbedarf:

Das Instrument gibt Hinweise auf das Vorliegen von Schmerzen, deren Intensität kann über die Skala jedoch nicht zuverlässig bestimmt werden, ebenso wenig die Schmerzlokalisierung und andere Schmerzqualitäten. Es ist eine niedrige Sensitivität für die Erfassung leichter Schmerzen anzunehmen. Die BESD-II Studie wirft weitere Fragen beispielsweise hinsichtlich der Ansprechbarkeit auf Veränderungen, aber auch der Wirksamkeit von Analgetika bei fortgeschrittener Demenz und fehlendem Placeboeffekt auf.

Literatur:

Deutsche Version unter

<https://www.dgss.org/die-gesellschaft/arbeitskreise/schmerz-und-alter/downloads/>

Basler HD, Hüger D, Kunz R et al. (2006) Beurteilung von Schmerz bei Demenz (BESD). Schmerz 20:519-526

Lukas A, Hagg-Grun U, Mayer B et al. (2019) Pain assessment in advanced dementia. Validity of the German PAINAD - a prospective double-blind randomised placebo-controlled trial. Pain 160(3):742-753. DOI: 10.1097/j.pain.0000000000001430

Schuler MS, Becker S, Kaspar R et al. (2007) Psychometric properties of the German "Pain Assessment in Advanced Dementia Scale" (PAINAD-G) in nursing home residents. J Am Med Dir Assoc 8:388-395

Smith TO, Harvey K (2022) Psychometric properties of pain measurements for people living with dementia: a COSMIN systematic review. Eur Geriatr Med 13:1029–1045. <https://doi.org/10.1007/s41999-022-00655-z>

Tab. 2.7.2: Schmerz

Name des Instrumentes:	Beobachtungsinstrument für das Schmerzassessment bei alten Menschen mit Demenz		
Kürzel:	BISAD		
Mittlerer Zeitaufwand: Entfällt für Patient*in / für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min			
Lernaufwand für Untersucher*in:			
☐ leicht	☒ mittel	☐ anspruchsvoll	
Durchführung durch: Ärzt*in, Pflegekraft, Ergotherapeut*in, Physiotherapeut*in			
Geeignet für Schweregrad:			
☐ leicht/Früherkennung	☒ mittel	☒ schwer	
Gütekriterien: Fischer (2012): Validierung der deutschen Version an n = 149 demenziell erkrankten (Stadium 6 und 7 nach Reisberg) Bewohner*innen in 27 Pflegeheimen, 83,9 ± 8,5 (52 – 103) J, 83,2% F. Konstruktvalidierung: BISAD-Score bei schmerzverursachender Erkrankung in Bewegung im Mittel um 1,9 Punkte (p = 0, 005) höher als ohne solche Erkrankungen. Schwach positive Korrelationen zwischen BISAD-Score und auf Schmerz hinweisender Mimik sowie verbal aggressiven Verhaltensweisen. Interne Konsistenz: Cronbachs Alpha für BISAD in Ruhe 0,65, in Bewegung 0,66.			
Begründung der Aufnahme in die Leitlinie: Alternative zur BESD.			
Limitationen / Forschungsbedarf: Das Instrument gibt Hinweise auf das Vorliegen von Schmerzen, deren Intensität kann über die Skala jedoch nicht zuverlässig bestimmt werden, ebenso wenig die Schmerzlokalisierung und andere Schmerzqualitäten. Es ist eine niedrige Sensitivität für die Erfassung leichter Schmerzen anzunehmen. Untersuchungen zur Interrater- und Retest-Reliabilität fehlen bislang. Die niedrige interne Konsistenz deuten darauf hin, dass die Skalenstruktur verbessert werden kann. Die Validierung bei Klinikpatienten sowie zu Hause lebenden älteren Menschen steht noch aus. Da das aktuelle Verhalten zu dem gewöhnlichen in Beziehung gesetzt werden soll, sind Vorinformationen notwendig.			
Literatur: Fischer T (2012): Schmerzeinschätzung bei Menschen mit schwerer Demenz. Das Beobachtungsinstrument für das Schmerzassessment bei alten Menschen mit schwerer Demenz (BISAD). Bern: Verlag Hans Huber Fischer T (2009) Entwicklung eines Instruments zum Assessment von Schmerzen bei alten Menschen mit schwerer Demenz (Dissertation an der Charité Berlin) https://refubium.fu-berlin.de/bitstream/handle/fub188/9531/Dissertation_Abgabefassung.pdf?sequence=1&isAllowed=y (Zugriff 04.06.2022)			

2.8 Ernährung / Mundgesundheit / Dysphagie

Ernährung

Mangelernährung zählt zu den geriatrischen Syndromen und ist bei geriatrischen Patient*innen weit verbreitet [Kiesswetter 2016, Cereda 2012, Pirlich 2006]. Ein schlechter Ernährungszustand ist im Alter fast immer multifaktoriell verursacht und geht mit gravierenden negativen Folgen für Krankheitsverlauf und Prognose einher (z. B. Abnahme der Leistungsfähigkeit, reduzierte Lebensqualität, vermehrte Komplikationen im stationären Bereich, erhöhte Mortalität), die beim älteren Menschen sehr viel häufiger auftreten als bei jüngeren mangelernährten Patient*innen [Agarwal 2013]. Mangelernährung kann durch diverse Interventionen auch im Alter erfolgreich therapiert werden [Volkert 2019, 2013]. Wegen der schwerwiegenden gesundheitlichen Konsequenzen muss als Ziel angesehen werden, bereits das Entstehen einer Mangelernährung zu vermeiden. Das Vorliegen eines erhöhten Risikos für Mangelernährung soll daher routinemäßig bei Aufnahme ins Krankenhaus erhoben werden und sollte die folgenden Aspekte berücksichtigen [Kondrup 2003]:

1. niedriger BMI (→ grobe Einschätzung des aktuellen Ernährungszustands)
2. (unbeabsichtigter) Gewichtsverlust
(→ negative Entwicklung des Ernährungszustands in der Vergangenheit)
3. geringe Essmenge
4. erhöhter Nährstoff-Bedarf
(→ bei 3. und 4. hohes Risiko für negative Entwicklung des Ernährungszustands in der Zukunft).

Das **Malnutrition Screening Tool (MST)** (Tabelle 2.8.1) stellt eine verkürzte Variante des Subjective Global Assessment (SGA) dar und enthält nur dessen Fragen mit der höchsten Sensitivität und Spezifität. Es ist sehr einfach und schnell durchzuführen, da nur ungewollter Gewichtsverlust und reduzierte Nahrungsaufnahme aufgrund eines verminderten Appetits abgefragt werden, Vorkenntnisse des durchführenden Personals sind nicht erforderlich. Bei kognitiv eingeschränkten Patient*innen ist eine Fremdanamnese möglich.

Das von der British Association for Parenteral and Enteral Nutrition (BAPEN) primär für den ambulanten Bereich entwickelte **Malnutrition Universal Screening Tool (MUST)** [Stratton 2004, BAPEN 2011] zeigte in einer Meta-Analyse [Cortez-Aguilar 2024] eine Sensitivität von 97% und Spezifität von 80% bei Krankenhaus-Patient*innen in der Erfassung von Mangelernährung gemäß der Kriterien der European Society for Clinical Nutrition and Metabolism [ESPEN].

Im Kontext des Assessments von geriatrischen Patient*innen ist besonders die validierte **Kurzform des Mini Nutritional Assessment (MNA-SF)**, bestehend aus 6 Fragen, Score 0-14 Punkte) (Tabelle 2.8.2) weit verbreitet [Power 2018a,b]. Sie wurde wie auch die **MNA-Langform**, aus der sie entwickelt wurde (**MNA-LF**, bestehend aus 18 Fragen, Score 0 – 30 Punkte), speziell für ältere Menschen konzipiert. Die internationale Verbreitung wurde anschließend von der Firma Nestlé gefördert, welche das Copyright für beide Versionen hält (www.mna-elderly.com). Die MNA-SF erfasst zusätzlich zu den Kern-Parametern (s. o.) Informationen über altersspezifische Risikofaktoren für eine

Mangelernährung (eingeschränkte Mobilität, beeinträchtigte Kognition) [Kaiser 2009], die aus dem geriatrischen Assessment übernommen werden können [Christner 2016]. Bei kognitiv eingeschränkten Patient*innen ist eine Fremdanamnese möglich.

Das **Nutritional Risk Screening (NRS 2002)** (Tabelle 2.8.3) soll Personen identifizieren, die ernährungsmedizinischer Unterstützung bedürfen [Kondrup 2003, Schütz 2005]. Der zweistufige Fragebogen besteht aus einem Vorscreening und einem Hauptscreening und berücksichtigt die besondere Risikosituation geriatrischer Patient*innen durch Einbeziehung des Schweregrads der Erkrankung und die automatische Höherstufung des Risikos für Menschen über 70 Jahre.

Die Übereinstimmung mit den Kriterien für Mangelernährung der Global Leadership Initiative on Malnutrition (GLIM) liegt (über Cohens kappa) für das MST bei 0,59, für das MUST bei 0,39 (einschließlich Risiko für Mangelernährung: 0,75) und für das MNA-SF bei 0,21 (einschließlich Risiko für Mangelernährung: 0,45) [van Dronkelaar 2023]. Sehr wichtig ist in jedem Fall, bei erkanntem Risiko eine detailliertere Erfassung und Abklärung der Ursachen mit einer angemessenen Maßnahmenplanung und anschließender Erfolgskontrolle durchzuführen [Volkert 2019, Volkert 2013]. Wird aktuell kein Risiko für eine Mangelernährung erkannt, sollte dies in regelmäßigen Abständen und bei jeder Veränderung der Versorgungssituation überprüft werden.

Tab. 2.8.1: Dimension: Ernährung

Name des Instrumentes:	Malnutrition Screening Tool		
Kürzel:	MST		
Mittlerer Zeitaufwand: Für Patient*in 1 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 3 min			
Lernaufwand für Untersucher*in:			
☒ leicht		☐ mittel	☐ anspruchsvoll
Durchführung durch: Pflegekraft, Ärzt*in, Ernährungsfachkraft, sonstiges Personal, Angehörige, Patient*in selbst			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung		☒ mittel	☐ schwer
Gütekriterien: Wurde für Personen jeden Alters entwickelt und für ältere Patient*innen im Krankenhaus-Setting validiert [Power 2018a].			
Begründung der Aufnahme in die Leitlinie: Instrument mit dem geringsten Zeitbedarf.			

Limitationen / Forschungsbedarf:

Berücksichtigt nur die beiden Risikofaktoren Verlust von Körpermasse und Appetit; im Alter sehr häufige Risikokonstellationen für das zukünftige Entstehen einer Mangelernährung (z. B. Krankheiten, kognitive Einschränkungen) werden nicht beachtet. Nicht durchführbar, wenn Gewichtsverlauf nicht ermittelbar. Deutsche Version nicht validiert.

Literatur:

Ferguson M, Capra S, Bauer J, Banks M (1999) Development of a valid and reliable malnutrition screening tool for adult acute hospital patients. Nutrition 15:458-464

Power L, Mullally D, Gibney ER. et al. A review of the validity of malnutrition screening tools used in older adults in community and healthcare settings - A MaNuEL study. Clin Nutr ESPEN 2018a; 24:1-13

Power L, de van der Schueren MA, Leij-Halfwerk S et al (2019) Development and application of a scoring system to rate malnutrition screening tools used in older adults in community and healthcare settings – a MaNuEL study. Clin Nutr 38(4):1807-1819. DOI: 10.1016/j.clnu.2018.07.022

Power L, Mullally D, Gibney ER et al. (2018a) A review of the validity of malnutrition screening tools used in older adults in community and healthcare settings - A MaNuEL study. Clin Nutr ESPEN 24:1-13

Power L, de van der Schueren MA, Leij-Halfwerk S et al. (2019) Development and application of a scoring system to rate malnutrition screening tools used in older adults in community and healthcare settings – A MaNuEL study. Clin Nutr 38(4):1807-1819. DOI: 10.1016/j.clnu.2018.07.022

Serón-Arbeloa C, Labarta-Monzón L, Puzo-Foncillas J (2022) Malnutrition screening and assessment. Nutrients 14, 2392. <https://doi.org/10.3390/nu14122392>

Tab. 2.8.2: Dimension: Ernährung

Name des Instrumentes:	Mini Nutritional Assessment – Kurzform	
Kürzel:	MNA-SF	
Mittlerer Zeitaufwand:		
Für Patient*in 5 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 7 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Pflegekraft, Ärzt*in, Ernährungsfachkraft, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien:		
Gute prädiktive Validität für Mortalität und andere klinische Endpunkte, jedoch Sensitivität gegenüber den Kriterien für Mangelernährung der Global Leadership Initiative on Malnutrition (GLIM) begrenzt.		
Kaluźniak-Szymanowska (2025): (n = 5614) Kohortenstudie. Sensitivität 69,5 % (Risiko für Mangelernährung gemäß GLIM), Spezifität 82,6%, positiver prädiktiver Wert 0,50, negativer		

prädiktiver Wert 0,91.

van Dronkelaar (2023): n = 356, Median 70 J, 46% F, multizentrische Studie, hospitalisierte Patient*innen mit einer Prävalenz von Mangelernährung gemäß GLIM-Kriterien von 42%. Sensitivität 32% (Risiko für Mangelernährung) bzw. 68 % (Mangelernährung), Spezifität 61% bzw. 98 %. Positiver prädiktiver Wert 0,90 bzw. 0,62, negativer prädiktiver Wert 0,67 bzw. 0,86.

Begründung der Aufnahme in die Leitlinie:

Zurzeit am breitesten eingeführtes Instrument zur Erfassung von Hinweisen auf Mangelernährung bei geriatrischen Krankenhauspatienten.

Limitationen / Forschungsbedarf:

Nicht durchführbar, wenn der BMI nicht erfasst werden kann, hier kann allerdings auf eine (weniger gut validierte) Version des MNA-SF zurückgegriffen werden, bei der stattdessen der Wadenumfang gemessen wird [Kaiser 2009, Power 2018a]. Ggf. minimales Training und Ausstattung für Bestimmung von Gewicht, Größe und/oder Wadenumfang erforderlich.

Literatur:

Deutsche Version auf der Homepage der Deutschen Gesellschaft für Ernährungsmedizin (DGEM) unter <https://www.dgem.de/screening>,

Ausfüllhinweise unter https://www.mna-elderly.com/forms/mna_guide_german_sf.pdf

Christner S, Ritt M, Volkert D et al. (2016) Evaluation of the nutritional status of older hospitalized geriatric patients: a comparative analysis of a Mini Nutritional Assessment (MNA) version and the Nutritional Risk Screening (NRS 2002). J Hum Nutr Diet 29:704-713

Cruz-Jentoft AJ, Volkert D (2025) Malnutrition in Older Adults. N Engl J Med 392(22):2244-2255. doi: 10.1056/NEJMr2412275
van Dronkelaar C, Tieland M, Cederholm T et al. (2023) Malnutrition Screening Tools are not sensitive enough to identify older hospital patients with malnutrition. Nutrients 15(24):5126. DOI:10.3390/nu15245126

Kaiser MJ, Bauer JM, Ramsch C et al. (2009) Validation of the Mini Nutritional Assessment short-form (MNA-SF): a practical tool for identification of nutritional status. J Nutr Health Aging 13:782-788

Kaluźniak-Szymanowska A, Deskur-Śmielecka E, Krzyżnińska-Siemaszkó R et al. (2025) Diagnostic Performance of Mini Nutritional Assessment-Short Form Against the Global Leadership Initiative on Malnutrition Criteria. Scientific Reports 15:45621. DOI:10.1038/s41598-025-30004-6

Power L, Mullally D, Gibney ER et al. (2018a) A review of the validity of malnutrition screening tools used in older adults in community and healthcare settings - A MaNuEL study. Clin Nutr ESPEN 24:1-13

Power L, de van der Schueren MA, Leij-Halfwerk S et al. (2019) Development and application of a scoring system to rate malnutrition screening tools used in older adults in community and healthcare settings – A MaNuEL study. Clin Nutr 38(4):1807-1819. DOI: 10.1016/j.clnu.2018.07.022

Rubenstein LZ, Harker JO, Salva A et al. (2001) Screening for undernutrition in geriatric practice: developing the short-form Mini-Nutritional Assessment (MNA-SF). J Gerontol A Biol Sci Med Sci 56:M366-372

Tab. 2.8.3: Dimension: Ernährung

Name des Instrumentes:	Nutritional Risk Screening	
Kürzel:	NRS 2002	

Mittlerer Zeitaufwand:
Für Patient*in 5 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 7 min

Lernaufwand für Untersucher*in:

☐ leicht	☒ mittel	☐ anspruchsvoll
------------------------------	---	-------------------------------------

Durchführung durch:
Pflegekraft, Ärzt*in, Ernährungsfachkraft, spezifisch geschultes Fachpersonal

Geeignet für Schweregrad:

☒ leicht/Früherkennung	☒ mittel	☒ schwer
---	---	---

Gütekriterien:
Wurde für Klinikpatienten jeden Alters entwickelt und für ältere Patient*innen im Krankenhaus-Setting validiert [Power 2018a]. Im Vergleich zur MNA-SF werden weniger (nämlich 66,0% der Patient*innen in geriatrischen Abteilungen) als von Mangelernährung bedroht eingestuft [Christner 2016].

Begründung der Aufnahme in die Leitlinie:
Erlaubt Anwendung bei stationären Patient*innen aller Fachabteilungen, gibt Handlungsempfehlungen.

Limitationen / Forschungsbedarf:
Nicht durchführbar, wenn BMI/Körpergewicht und Gewichtsverlauf oder Nahrungszufuhr nicht erfasst werden können. Ggf. minimales Training und Ausstattung für Bestimmung von Gewicht und Größe erforderlich. Gewichtsverlust, Abschätzung einer prozentuell geringeren Nahrungsaufnahme im Vergleich zum Bedarf sowie Krankheitsschwere ggf. nicht einfach zu ermitteln und zu kategorisieren (Kategorien der Krankheitsschwere sind nicht auf geriatrische Patient*innen abgestimmt), Erfahrung/Schulung notwendig. Spezifität des Instrumentes leidet für ältere Personen durch die automatische Höherstufung von Patient*innen über 70 Jahre.

Literatur:
Deutsche Version auf der Homepage der Deutschen Gesellschaft für Ernährungsmedizin (DGEM) <https://www.dgem.de/instrumente-zur-erfassung-des-ernahrungszustands> (Zugriff 29.12.2023)
Christner S, Ritt M, Volkert D et al. (2016) Evaluation of the nutritional status of older hospitalized geriatric patients: a comparative analysis of a Mini Nutritional Assessment (MNA) version and the Nutritional Risk Screening (NRS 2002). J Hum Nutr Diet 29:704-713
Kondrup J, Rasmussen HH, Hamberg O et al. (2003) Nutritional risk screening (NRS 2002): a new method based on an analysis of controlled clinical trials. Clin Nutr 22:321-336
Power L, Mullally D, Gibney ER et al. (2018a) A review of the validity of malnutrition screening tools used in older adults in community and healthcare settings - A MaNuEL study. Clin Nutr ESPEN 24:1-13
Power L, de van der Schueren MA, Leij-Halfwerk S et al. (2019) Development and application of a

scoring system to rate malnutrition screening tools used in older adults in community and healthcare settings – A MaNuEL study. Clin Nutr 38(4):1807-1819. DOI: 10.1016/j.clnu.2018.07.022

Power L, de van der Schueren MA, Leij-Halfwerk S et al (2019) Development and application of a scoring system to rate malnutrition screening tools used in older adults in community and healthcare settings – a MaNuEL study. Clin Nutr 38(4):1807-1819. DOI: 10.1016/j.clnu.2018.07.022

Power L, Mullally D, Gibney ER et al. (2018a) A review of the validity of malnutrition screening tools used in older adults in community and healthcare settings - A MaNuEL study. Clin Nutr ESPEN 24:1-13

Power L, de van der Schueren MA, Leij-Halfwerk S et al. (2019) Development and application of a scoring system to rate malnutrition screening tools used in older adults in community and healthcare settings – A MaNuEL study. Clin Nutr 38(4):1807-1819. DOI: 10.1016/j.clnu.2018.07.022

Mundgesundheit

Der orale Gesundheitszustand kann sich auf das Sprechen, häufiger jedoch auf die Ernährung älterer Menschen und ihre Lebensqualität auswirken. Chronische Entzündungen, z. B. Parodontitis, stellen darüber hinaus einen Risikofaktor für vaskuläre Komplikationen dar. Aus diesen Gründen sollte die Mundgesundheit in das geriatrische Assessment einbezogen werden.

Zur Selbstauskunft bei kognitiv gesunden Patient*innen wird international der **Geriatric (auch „General“) Oral Health Assessment Index** (GOHAI) [Atchison 1990] eingesetzt, der fünfstufig die Häufigkeit des Auftretens von Beeinträchtigungen mit 12 von der Mundgesundheit abhängigen Items erfasst. Eine validierte deutsche Übersetzung [Hassel 2008] ist verfügbar.

Das **Oral Health Assessment Tool** (OHAT) [Chalmers 2004] bewertet acht Domänen der Mundgesundheit (Lippen, Zunge, Zahnfleisch, Speichel, Zähne, Zahnprothesen, Sauberkeit des Mundes und Zahnschmerzen) dreistufig. Die deutsche Übersetzung zeigte bei Anwendung durch trainiertes Pflegefachpersonal Werte von 87% bzw. 86% für Spezifität und Sensitivität im Vergleich zum zahnärztlichen Urteil als Goldstandard [Klotz 2020].

Dysphagie (Aspirationsrisiko)

Das Risiko für Dysphagien (Schluckstörungen) steigt mit zunehmendem Lebensalter an: So sind circa 50% der Bewohner*innen von Pflegeheimen und bis zu 70% der Patienten in geriatrischen Kliniken betroffen [Lin 2002] aber auch 14% der selbständig lebenden Bevölkerung über 65 Jahre [Turley 2009]. Alternsphysiologische Veränderungen, die zu einer Verminderung der Sicherheit und Effizienz des Schluckvorganges führen, werden unter dem Begriff Presbyphagie zusammengefasst; bei Verlust der Kompensationsfähigkeit resultiert eine Presbydysphagie. Daneben spielen altersassoziierte Erkrankungen, vorrangig aus dem Fachgebiet der Neurologie,

ursächlich eine entscheidende Rolle. Als Beispiele sind Schlaganfall, Demenz, Parkinson oder amyotrophe Lateralsklerose anzuführen. Klinisch relevante Folgen von Schluckstörungen sind Mangelernährung, Dehydratation sowie Aspiration mit hohem Pneumonierisiko [Serra-Prat 2012], daneben können Auswirkungen auf die soziale Interaktion mit Rückzug und Isolation resultieren. Ziel des Dysphagie-/Aspirations-Screenings ist es, einfach, schnell und möglichst zuverlässig aspirationsgefährdete Patient*innen zu identifizieren, um prophylaktische Maßnahmen und weiterführende Diagnostik einleiten zu können.

Folgende Symptome sind eigen- und/oder fremdanamnestisch zu erheben:

Häufiges Verschlucken, veränderte Stimmqualität, Erstickungsanfälle oder Husten nach Aufnahme von Nahrung oder Flüssigkeiten, verminderte Essens- und Trinkmengen, „Steckenbleiben“ von Speisen oder Getränken, ungewollter Gewichtsverlust, veränderte Kopfhaltung während des Schluckens (z. B. Anteflexion), wiederholte unklare Fieberschübe und Pneumonie (als Hinweis auf Aspiration).

Bei Verdacht auf eine Dysphagie bzw. ein erhöhtes Aspirationsrisiko sollen standardisierte Verfahren zum Einsatz kommen, die eine entsprechende Sensitivität und Spezifität bieten, wie das **Standardised Swallowing Assessment (SSA)** (Tabelle 2.8.4) oder die **Prädiktoren der Aspiration nach Daniels** (Tabelle 2.8.5). Anzumerken ist, dass praktisch alle Screening- und Assessmentinstrumente für die Zielgruppe Schlaganfall entwickelt wurden. Die S1-Leitlinie Neurogene Dysphagie [AWMF Reg.-Nr. 030-111] macht auf das Fehlen direkter Vergleichsstudien aufmerksam und empfiehlt, in die Auswahl des „jeweils optimalen“ Testverfahrens weitere Faktoren, z. B. Patienten-Charakteristika, einzubeziehen.

Die interprofessionelle Arbeitsgruppe „Dysphagie“ der Deutschen Gesellschaft für Geriatrie (DGG) hat das **Dysphagie Screening-Tool Geriatrie (DSTG)** (Tabelle 2.8.6) vorgeschlagen, das auf Basis einer selektiven Literaturrecherche und subjektiver Erfahrungen interprofessionell konsentiert wurde. Das Instrument soll Hinweise auf eine oropharyngeale Dysphagie bei geriatrischen Patient*innen erfassen. Es gliedert sich in drei Teile: 1. Vigilanz und Sitzposition, 2. Speichelschluck, Zungenbeweglichkeit und Hustenstoß sowie 3. Wasserschlucktest. Das Instrument kann durch geschultes medizinisches Personal angewendet werden und ist einfach in den Arbeitsalltag zu integrieren. Eine klare Anleitung und eine Validierung gegen endoskopische Schluckdiagnostik (FEES) liegen vor.

Mehr-Konsistenzen-Tests, die neben Wasser auch andere Nahrungskonsistenzen evaluieren, erlauben eine genauere Graduierung der Dysphagie und damit differenziertere Empfehlungen für die Kostform sowie das weitere diagnostische und therapeutische Vorgehen.

Bei Vorliegen einer Dysphagie ist eine klinische Schluckuntersuchung (KSU) durch Sprachtherapeut*innen/Logopäd*innen zu fordern. Für die Therapieplanung sind

Verfahren notwendig, die eine genauere Betrachtung der Vorgänge während des Schluckens erlauben (Videofluoroskopie - Videofluoroscopic Swallowing Study, VFSS und Videoendoskopie des Schluckens - Flexible Endoscopic Evaluation of Swallowing, FEES).

Tab. 2.8.4: Ernährung/Dysphagie

Name des Instrumentes:	Standardised Swallowing Assessment	
Kürzel:	SSA	
Mittlerer Zeitaufwand:		
Für Patient*in 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 15 min		
Lernaufwand für Untersucher*in:		
☐ leicht	☒ mittel	☐ anspruchsvoll
Durchführung durch:		
Pflegefachkraft, Ärzt*in, Logopäd*in		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien:		
Wang (2025): n = 466 Personen über 60 J in China, die in den letzten 12 Monaten nicht im Krankenhaus aufgenommen worden waren; im Mittel 62 J, 61% F. Retest-Reliabilität der chinesischen Version (n = 79) >0,8, Interrater-Reliabilität (n = 149) >0,9 (Korrelation nach Pearson), AUC 0,80.		
Positiver prädiktiver Wert für das Vorliegen einer Dysphagie 0,84, negativer prädiktiver Wert 0,89. Sensitivität 97% und Spezifität 90% bezüglich des Erkennens einer Dysphagie.		
Begründung der Aufnahme in die Leitlinie:		
Hohe Relevanz des frühzeitigen Erkennens einer Dysphagie.		
Limitationen / Forschungsbedarf:		
Der SSA ist ein geeignetes klinisches Instrument, um eine mögliche Schluckstörung zu erfassen. Für eine Therapieplanung reicht das Ergebnis nicht aus.		
Literatur:		
Dziewas R, Pflug C et al., Deutsche Gesellschaft für Neurologie (2020) S1-Leitlinie Neurogene Dysphagie. AWMF-Register-Nr. 030-111		
Perry L (2001) Screening swallowing function of patients with acute stroke. Part one: identification, implementation and initial evaluation of a screening tool for use by nurses. J Clin Nurs 10:463–473		
Wang J, Chen C, Qin Y et al. (2025) Reliability and validity of the Standardized swallowing assessment among community-dwelling older adults in China. Ann Med 57(1):2548980. DOI: 10.1080/07853890.2025.2548980		
Westergren A (2006) Detection of eating difficulties after stroke: a systematic review. Int Nurs Rev		

53:143–149
Zhang L, Hou R, Liu L et al. (2025) Evaluation of the performance of screening tools for dysphagia in older adults: A diagnostic meta-analysis. Geriatr Nurs 61:629-641. DOI: 10.1016/j.gerinurse.2024.12.039

Tab. 2.8.5: Ernährung/Dysphagie

Name des Instrumentes:	Prädiktoren der Aspiration nach Daniels		
Kürzel:	–		
Mittlerer Zeitaufwand:			
Für Patient*in 10 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 15 min			
Lernaufwand für Untersucher*in:			
☐ leicht	☒ mittel	☐ anspruchsvoll	
Durchführung durch:			
Logopäd*in, Ärzt*in			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien:			
Das Verfahren hat eine Sensitivität von 93% und eine Spezifität von 67% (Referenzstandard: Videofluoroskopie).			
Begründung der Aufnahme in die Leitlinie:			
Hohe Relevanz des frühzeitigen Erkennens einer Dysphagie.			
Limitationen / Forschungsbedarf:			
Der Test nach Daniels ist ein geeignetes klinisches Instrument, um eine mögliche Schluckstörung zu erfassen. Für eine Therapieplanung reicht das Ergebnis nicht aus.			
Literatur:			
<u>Daniels SK, McAdam CP, Brailey K et al. (1997) Clinical assessment of swallowing and prediction of dysphagia severity. Am J Speech Lang Pathol 6:17–24</u>			
<u>Daniels SK, Ballo LA, Mahoney MC et al. (2000) Clinical predictors of dysphagia and aspiration risk: outcome measures in acute stroke patients. Arch Phys Med Rehabil 81:1030–1033</u>			
<u>Dziewas R, Pflug C et al., Deutsche Gesellschaft für Neurologie (2020) S1-Leitlinie Neurogene Dysphagie. AWMF-Register-Nr. 030-111</u>			

Tab. 2.8.6: Ernährung/Dysphagie

Name des Instrumentes: Kürzel:	Dysphagie Screening-Tool Geriatrie DSTG		
Mittlerer Zeitaufwand: Für Patient*in 13 min / für Untersucher*in (inkl. Vor- u. Nachbereitung) 18 min			
Lernaufwand für Untersucher*in:			
☐ leicht	☒ mittel	☐ anspruchsvoll	
Durchführung durch: Pflegekraft, Ärzt*in, Logopäd*in			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien: Thiem (2023): n = 53, stationäre Patient*innen in 5 Geriatrien, 86 ± 6 J, 57% F. FEES und DGST in zufälliger Reihenfolge. 20 der Studienteilnehmenden wiesen gemäß FEES eine Dysphagie auf, davon erkannte der DSTG 12, Sensitivität 60%, Spezifität 88%, positiver prädiktiver Wert 0,75, negativer prädiktiver Wert 0,78, AUC in der ROC 0,77.			
Begründung der Aufnahme in die Leitlinie: Speziell für den Einsatz in der Geriatrie konzipiertes Instrument zur Dysphagie-Erkennung.			
Limitationen / Forschungsbedarf: Für eine Therapieplanung reicht das Ergebnis nicht aus.			
Literatur: Dziewas R, Pflug C et al., Deutsche Gesellschaft für Neurologie (2020) S1-Leitlinie Neurogene Dysphagie. AWMF-Register-Nr. 030-111 <u>Jäger M, Thiem U, Stege H (2020) Entwicklung eines neuen Screeninginstruments zum Screening auf Dysphagie bei geriatrischen Patienten: das Dysphagie Screening-Tool Geriatrie. Z Gerontol Geriatr. 53(3):239-244. DOI: 10.1007/s00391-020-01712-7</u> Thiem U, Jäger M, Stege H, Wirth R (2023) Diagnostic accuracy of the 'Dysphagia Screening Tool for Geriatric Patients' (DSTG) compared to Flexible Endoscopic Evaluation of Swallowing (FEES) for assessing dysphagia in hospitalized geriatric patients – a diagnostic study. BMC Geriatrics 23, 856. DOI: 10.1186/s12877-023-04516-7			

2.9 Kontinenz

Kontinenz erfordert ein so komplexes Zusammenspiel auf neuronaler und motorischer Ebene, dass zahlreiche Erkrankungen, Traumata und Pharmaka sie beeinträchtigen können. Während Stuhlinkontinenz in Item 9 des Barthel-Index erfasst wird, werden

hinsichtlich der Harninkontinenz in der Operationalisierung nach Hamburger Manual Betroffene, die es schaffen, ein Einnässen von Bett und Kleidung zu verhindern, nicht erfasst (in der Operationalisierung nach Barthel plus erhalten sie die Bewertung 10b, ohne Harninkontinenz 10a). Eine Einstiegsfrage (z. B. „Verlieren Sie ungewollt Urin oder Stuhl?“) soll Teil der ärztlichen Aufnahmeanamnese sein, sofern diese nicht die Erhebung des Barthel-Index inkludiert. Für die sich ggf. anschließende Befragung kommen z. B. die vier Fragen umfassende International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form (ICIQ-UI SF, www.iciq.net) [Ressler 2010] in Betracht. Näheres, Alternativen und Empfehlungen für das weitere diagnostische und therapeutische Vorgehen sind der Leitlinie „Harninkontinenz bei geriatrischen Patienten“ [AWMF 084-001] zu entnehmen.

2.10 Schlaf

Mit dem **Essener Fragebogen Alter und Schläfrigkeit (EFAS)** (Tabelle 2.10.1) steht das bisher einzige, speziell für geriatrische Patient*innen entwickelte valide Messinstrument zur Erfassung von Tagesschläfrigkeit zur Verfügung. Der Fragebogen wurde als Fremdbeurteilungsinstrument entwickelt. Pflege- oder Betreuungspersonen beobachten den Patient*innen über einen Zeitraum von einer Woche und beurteilen die beobachtete Häufigkeit des Schlafens am Tage zu vier vorgegebenen Zeiten. Zudem wird die Schlaftiefe qualitativ beurteilt. Für die Gesamtbeurteilung wurde die jeweils höchste Bewertung eines von vier Items in den jeweiligen Kategorien Häufigkeit und Schweregrad verwendet und das Produkt gebildet. Die Interrater- und die Retest-Reliabilität für den Gesamtscore wurde durch Berechnung der Korrelationskoeffizienten und Cohens κ bestimmt. Für eine Subgruppe von Patient*innen mit erhaltener Selbstbeurteilungsfähigkeit und ohne relevante Hirnleistungsstörung erfolgte eine externe Validierung. Hierzu diente die Frage nach der selbst eingeschätzten Einschlafwahrscheinlichkeit aus dem Schlaffragebogen nach Siegrist. Die fünfstellige Ratingskala wurde dichotomisiert in die beiden Antwortgruppen „nie/selten“ und „gelegentlich/häufig/sehr häufig“ und mittels ROC-Analyse mit dem Produktscore (Häufigkeit x Schweregrad, 0–12) der Fremdbeurteilungsskala verglichen. Zudem erfolgte eine externe Validierung mittels Pupillographie, bei der sich akzeptable Werte fanden [Frohnhofer 2013].

Tab. 2.10.1: Dimension: Schlaf

Name des Instrumentes:	Essener Fragebogen Alter und Schläfrigkeit	
Kürzel:	EFAS	
Mittlerer Zeitaufwand:	Entfällt für Patient*in / für Untersucher*in (inkl. Vor- u. Nachbereitung) 10 min	
Lernaufwand für Untersucher*in:		
	x leicht	O mittel
		O anspruchsvoll

Durchführung durch: Arzt*in, Psycholog*in, spezifisch geschultes Fachpersonal		
Geeignet für Schweregrad:		
x leicht/Früherkennung	x mittel	x schwer
Gütekriterien: Mit einer Sensitivität von 81 % und einer Spezifität von 80 % erfasst ein EFAS-Wert > 2 im Gesamtscore eine zumindest als moderat empfundene Tagesschläfrigkeit.		
Begründung der Aufnahme in die Leitlinie: Tagesschläfrigkeit kann über die damit verbundene Bewegungsarmut einen Risikofaktor für Sarkopenie und weitere Erkrankungen darstellen.		
Limitationen / Forschungsbedarf: Der EFAS wurde als Fremdbeurteilungsskala in der stationären Akutgeriatrie validiert. Gütekriterien für Patienten in Rehabilitationskliniken, Heimbewohner oder ambulante Patienten liegen nicht vor.		
Literatur: Frohnhofer H, Bibl M, Nickel B, Popp R (2010) Der Essener Fragebogen Alter und Schläfrigkeit (EFAS) – ein neues Assessmentinstrument zur Messung von Tagesschläfrigkeit bei alten Menschen. Euro J Ger 12:84-90 Frohnhofer H, Fulda S, Frohnhofer K, Popp R (2013) Validation of the Essener Questionnaire of Age and Sleepiness in the elderly using pupillometry. Adv Exp Med Biol 755:125-132		

2.11 Substanzmissbrauch/Sucht

Der Missbrauch von Alkohol und anderen Substanzen spielt bei geriatrischen Patient*innen sowohl als in das Alter mitgenommenes als auch neu aufgetretenes Problem eine wichtige Rolle [Wolter 2011].

Der einfache **CAGE**-Fragebogen (Tabelle 2.11.1) aus 4 Fragen, die mit Ja oder Nein zu beantworten sind, ist Ausgangspunkt vieler Alkohol-Screeninginstrumente [Wolter 2011]: Das Akronym CAGE steht für **C**ut down on drinking - **A**nnoyed by criticism - **G**uilty feeling - **E**ye opener. Während im mittleren Lebensalter der Cut-off bei 1/2 positiven Antworten liegt, wird für ältere Menschen zur Verbesserung der Sensibilität die Absenkung vorgeschlagen: dann steht schon bei nur einer positiven Antwort der Verdacht auf ein Alkoholproblem im Raum, dem weiter nachgegangen werden sollte [Rumpf 2006].

Der **Alcohol Use Disorder Identification Test (AUDIT)** gilt als internationaler Goldstandard für das Screening von riskantem Alkoholkonsum, schädlichem Alkoholgebrauch oder Alkoholabhängigkeit. Er besteht aus zehn Fragen zu Trinkmengen und -gewohnheiten sowie möglichen Folgen eines übermäßigen Alkoholkonsums. Häufig wird die Kurzversion **AUDIT-C** (C steht für consumption) (Tabelle 2.11.2) mit den ersten drei Fragen zu Trinkmengen und -gewohnheiten verwendet, diese Kurzfassung erscheint – insbesondere im Alter [Berk 2008] – kaum schlechter als das vollständige

Instrument mit zehn Fragen [Kriston 2008, Reinert 2007]. Beide Versionen werden in der S3-Leitlinie „Screening, Diagnose und Behandlung alkoholbezogener Störungen“ [AWMF-Register-Nr. 076-001] empfohlen. Vor der Anwendung ist es wichtig, sich mit der Definition von „Standarddrink“ vertraut zu machen: Diese Maßeinheit enthält 12,0-13,2 g Reinalkohol und entspricht 1 Glas/Dose Bier (0,33 l) oder 1 Glas Wein/Sekt (0,125 l) oder 1 Glas Korn, Rum, Schnaps, Weinbrand, Whisky o.ä. (0,04 l, „doppelter Schnaps“).

Der **SMAST-G** (Tabelle 2.11.3) ist eine altersspezifische Modifikation des Short Michigan Alcohol Screening Test. Er besteht aus 10 Fragen zu Trinkgewohnheiten und den Auswirkungen des Alkoholkonsums, die mit Ja oder Nein zu beantworten sind. Moore et al. schlagen vor, CAGE und SMAST-G parallel zu verwenden, da der Zeitaufwand (v. a. für den CAGE-Fragebogen) gering, die Überlappung zwischen beiden Instrumenten aber nicht vollständig ist.

Assessment-Instrumente bzw. Rating Skalen zum Einsatz bei V. a. Missbrauch oder Abhängigkeit von Schlaf- und Beruhigungsmitteln (Benzodiazepine und vergleichbare Substanzen) werden im klinischen Alltag bisher kaum verwendet [Wolter 2011]. Der (Lippstädter) **Benzo-Check** (Tabelle 2.11.4) geht davon aus, dass Personen mit problematischem Benzodiazepin-Konsum sich in ihrem Selbstverständnis i. d. R. nicht als abhängig begreifen. Die klinischen Auffälligkeiten werden deshalb nicht als Symptome einer Suchterkrankung etikettiert. Die Verwendung der Begriffe „Sucht“ oder „Abhängigkeit“ wird vermieden, vielmehr werden Veränderungen, die die Betroffenen an sich bemerken bzw. unter denen sie leiden, als mögliche Folgen der Langzeiteinnahme deklariert. 5 der 12 Fragen zielen auf typische Symptome der Abhängigkeit, 7 weitere Fragen sind unspezifischer. Bei den 5 suchtspezifischen Fragen können je nach Ausprägungsgrad bis zu je 8 Punkte erreicht werden, bei den anderen Fragen bis zu je 4. Bei einer Gesamtpunktzahl zwischen 0 und 12 liegen „keine typischen, sicheren Folgen“ der Langzeiteinnahme vor, bei 13 bis 24 Punkten sind die Beschwerden „wahrscheinlich“ durch die Langzeiteinnahme bedingt, ab 25 Punkten besteht „hohe Wahrscheinlichkeit“.

Tab. 2.11.1: Dimension: Substanzmissbrauch und Sucht (Alkohol)

Name des Instrumentes:	CAGE-Fragebogen	
Kürzel:	CAGE	
Mittlerer Zeitaufwand:		
Für Patient*in 1 min/ für Untersucher*in (inkl. Vor- u. Nachbereitung) 1 min		
Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch:		
Ärzt*in, Pflegekraft, Ergotherapeut*in, Psycholog*in		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer

Gütekriterien:

Malet et al. 2009:
Cut-off 1/2: Sensitivität (m/w) 90%/100% → Spezifität (m/w) 79%/97%
Ryou et al. 2012:
Cut-off 2/3: Sensitivität 87% → Spezifität 63%
Cut-off 1/2: Sensitivität 100% → Spezifität 32%

Begründung der Aufnahme in die Leitlinie:

Instrument zur Ersteinschätzung auf Vorliegen eines Alkoholproblems.

Limitationen / Forschungsbedarf:

Der CAGE-Fragebogen zeigt Schwächen bei älteren Menschen mit psychiatrischer Komorbidität.

Literatur:

Deutsche Version unter <https://alkoholhilfe.at/wp-content/uploads/2018/06/INFO-Test-CAGE.pdf>
Berks J, McCormick R (2008) Screening for alcohol misuse in elderly primary care patients: a systematic literature review. Int Psychogeriatr 20:1090-1103
Malet L, Brousse G, Llorca PM (2009) CAGE unsound for alcohol misuse in seniors. Int J Geriatr Psychiatry 24:434-435
Moore AA, Seeman T, Morgenstern H et al. (2002) Are there differences between older persons who screen positive on the CAGE questionnaire and the Short Michigan Alcoholism Screening Test-Geriatric Version? J Am Geriatr Soc 50:858-862
O'Connell H, Chin AV, Hamilton F et al. (2004) A systematic review of the utility of self-report alcohol screening instruments in the elderly. Int J Geriatr Psychiatry 19:1074-1086
Rumpf HJ (2006) Diagnostik von alkoholbezogenen Störungen im Alter. Z Gerontopsychol psychiatr 19:201-206
Ryou YI, Kim JS, Jung JG et al. (2012) Usefulness of alcohol-screening instruments in detecting problem drinking among elderly male drinkers. Korean J Fam Med 33:126-133
St John PD, Montgomery PR, Tyas SL (2009) Alcohol misuse, gender and depressive symptoms in community-dwelling seniors. Int J Geriatr Psychiatry. 24:369-375
Taylor C, Jones KA, Denning T (2014) Detecting alcohol problems in older adults: can we do better? Int Psychogeriatr. 26:1755-1766

Tab. 2.11.2: Dimension: Substanzmissbrauch/Sucht

Name des Instrumentes:	Alcohol Use Disorder Identification Test – consumption		
Kürzel:	AUDIT-C		
Mittlerer Zeitaufwand:			
Für Patient*in 3 min/ für Untersucher*in (inkl. Vor- u. Nachbereitung) 4 min			
Lernaufwand für Untersucher*in:			
☒ leicht		☐ mittel	☐ anspruchsvoll

Durchführung durch: Ärzt*in, Pflegekraft, Ergotherapeut*in, Psycholog*in		
Geeignet für Schweregrad:		
x leicht/Früherkennung	x mittel	O schwer
Gütekriterien: AUC (95% Konfidenzintervall) 0,935 (0,914-0,957). Bei Cut-off von 3/4 liegt die Sensitivität bei 94%, die Spezifität bei 80% für das Erkennen von schädlichem Alkoholkonsum (Aalto 2011). Für Frauen wird ein Cut-off von 2/3 genannt.		
Begründung der Aufnahme in die Leitlinie: Auch bei älteren Personen, stationär und ambulant validierter Fragebogen, empfohlen in der S3-Leitlinie Screening, Diagnose und Behandlung alkoholbezogener Störungen.		
Limitationen / Forschungsbedarf: Definition von „Standarddrink“ muss beachtet werden.		
Literatur: Aalto M, Alho H, Halme JT, Seppä K (2011) The Alcohol Use Disorders Identification Test (AUDIT) and its derivatives in screening for heavy drinking among the elderly. Int J Geriatr Psychiatry 26:881-885 Berks J, McCormick R (2008) Screening for alcohol misuse in elderly primary care patients: a systematic literature review. Int Psychogeriatr 20:1090-1103 Berner MM, Kriston L, Bentele M, Härter M (2007) The alcohol use disorders identification test for detecting at-risk drinking: a systematic review and meta-analysis. J Stud Alcohol Drugs 68:461-473 Bisschop JM, de Jonge HJM, Brunsveld-Reinders AH et al. (2025) Screening instruments to detect problematic alcohol use among adults in hospitals and their diagnostic test accuracy: A systematic review. Kriston L, Hölzel L, Weiser AK et al. (2008) Meta-analysis: Are 3 questions enough to detect unhealthy alcohol use? Ann Intern Med 149:879-888 Reinert DF, Allen JP (2007): The alcohol use disorders identification test: an update of research findings. Alcohol Clin Exp Res 31:185-199 S3-Leitlinie "Screening, Diagnose und Behandlung alkoholbezogener Störungen". AWMF-Register-Nr. 076-001. Februar 2021 WHO (2001) The Alcohol Use Disorders Identification Test. Guidelines for Use in Primary Care. 2nd Edition. World Health Organization, Department of Mental Health and Substance Dependence. http://www.paho.org/English/DD/PUB/AuditBro-3.pdf		

Tab. 2.11.3: Dimension: Substanzmissbrauch/Sucht

Name des Instrumentes:	Short Michigan Alcoholism Screening Test Geriatric Version
Kürzel:	SMAST-G
Mittlerer Zeitaufwand: Für Patient*in 4 min/ für Untersucher*in (inkl. Vor- u. Nachbereitung) 56 min	

Lernaufwand für Untersucher*in:		
☒ leicht	☐ mittel	☐ anspruchsvoll
Durchführung durch: Ärzt*in, Pflegekraft, Ergotherapeut*in, Psycholog*in		
Geeignet für Schweregrad:		
☒ leicht/Früherkennung	☒ mittel	☐ schwer
Gütekriterien: Sensitivität 52-85% → Spezifität 93-97% → bei cut-off 2/3 [Naegle 2008] Sensitivität 75% → Spezifität 69% → bei cut-off 2/3 [Johnson-Greene 2009] Sensitivität 78% → Spezifität 58% → bei cut-off 2/3 [Ryou 2012]		
Begründung der Aufnahme in die Leitlinie: Speziell für ältere Personen konzipiertes Instrument.		
Limitationen / Forschungsbedarf: keine aktuellen wiss. Veröffentlichungen über den Einsatz im deutschen Sprachraum		
Literatur: Deutsche Version unter https://www.alterundsucht.ch/aerzteschaft/alkohol/screening.html (Zugriff 20.11.2025) Berks J, McCormick R (2008) Screening for alcohol misuse in elderly primary care patients: a systematic literature review. Int Psychogeriatr 20:1090-1103 Beullens J, Aertgeerts B (2004) Screening for alcohol abuse and dependence in older people using DSM criteria: a review. Aging Ment Health 8:76-82 Blow FC, Gillespie BW, Barry KL et al. (1998) Brief screening for alcohol problems in theelderly populations using the Short Michigan Alcoholism Screening Test-Geriatric Version (SMAST-G). Alcohol Clin Exp Res 22(Suppl):131 Johnson-Greene D, McCaul ME, Roger P (2009) Screening for hazardous drinking using the Michigan Alcohol Screening Test-Geriatric Version (MAST-G) in elderly persons with acute cerebrovascular accidents. Alcohol Clin Exp Res 33:1555-1561 Moore AA, Seeman T, Morgenstern H et al. (2002) Are there differences between older persons who screen positive on the CAGE questionnaire and the Short Michigan Alcoholism Screening Test-Geriatric Version? J Am Geriatr Soc 50:858-862 Naegle MA (2008) Screening for alcohol use and misuse in older adults: using the Short Michigan Alcoholism Screening Test - Geriatric Version. Am J Nurs 108:50-58 O'Connell H, Chin AV, Hamilton F et al. (2004) A systematic review of the utility of self-report alcohol screening instruments in the elderly. Int J Geriatr Psychiatry 19:1074-1086 Rumpf HJ (2006) Diagnostik von alkoholbezogenen Störungen im Alter. Z Gerontopsychol psychiatr 19:201-206 Ryou YI, Kim JS, Jung JG et al. (2012) Usefulness of alcohol-screening instruments in detecting problem drinking among elderly male drinkers. Korean J Fam Med 33:126-133		

Tab. 2.11.4: Dimension: Substanzmissbrauch/Sucht

Name des Instrumentes: Kürzel:	Benzo-Check —		
Mittlerer Zeitaufwand: Für Patient*in 10 min/ für Untersucher*in (inkl. Vor- u. Nachbereitung) 15 min			
Lernaufwand für Untersucher*in:			
☒ leicht	☐ mittel	☐ anspruchsvoll	
Durchführung durch: Ärzt*in, Pflegekraft, Ergotherapeut*in, Psycholog*in; Selbstausfüllen möglich			
Geeignet für Schweregrad:			
☒ leicht/Früherkennung	☒ mittel	☐ schwer	
Gütekriterien: Trennt mit hoher Zuverlässigkeit zwischen Benzodiazepin-Entzugspatienten und psychiatrischen Tagesklinikpatienten (Evaluationsstudie Holzbach 2019)			
Begründung der Aufnahme in die Leitlinie: Benzodiazepin-Abhängigkeit vermindert die Lebensqualität und erhöht das Demenzrisiko.			
Limitationen / Forschungsbedarf: Weitere Evaluationsstudien erforderlich, insbesondere für geriatrische Patienten. Bei kognitiver Beeinträchtigung je nach Schweregrad nur bedingt bis gar nicht einsetzbar.			
Literatur: Holzbach R (2017) Medikamentenabhängigkeit im Alter – Handlungsleitlinien zum pflegerischen Umgang. In: Hoff T, Kuhn U, Kuhn S, Isfort M (Hrsg.): Sucht im Alter – Maßnahmen und Konzepte für die Pflege. Springer, Berlin. 39-54 Holzbach R, Hunold P, Konert F (2021) Evaluation des Lippstädter Benzo-Checks als Screening-Instrument für Nebenwirkungen der Langzeiteinnahme von Benzodiazepinen und Z-Substanzen. Fortschr Neurol Psychiatr. DOI: 10.1055/a-1510-8425 Hunold P (2014) Psychiatrische Beeinträchtigungen durch den Langzeitkonsum von Benzodiazepinen. Dissertation Psychiatrische Universitätsklinik Hamburg. https://ediss.sub.uni-hamburg.de/handle/ediss/5709			

3 Literatur

Quellen, die sich einzelnen tabellarisch beschriebenen Testinstrumenten zuordnen lassen, werden dort genannt.

Agarwal E, Miller M, Yaxley A, Isenring E (2013) Malnutrition in the elderly: a narrative review. *Maturitas* 76:296-302 DOI: 10.1016/j.maturitas.2013.07.013

Arbeitsgruppe Geriatrisches Assessment (AGAST) (1995) Geriatrisches Basisassessment. Handlungsanleitungen für die Praxis. Schriftenreihe Geriatrie Praxis. MMV Medizin Verlag München. ISBN 3-8208-1273-3

Atchison KA, Dolan TA (1990) Development of the Geriatric Oral Health Assessment Index. *J Dent Educ* 54:680-687

BAPEN, Malnutrition Advisory Group. (2011). The 'MUST' explanatory booklet. A Guide to the 'Malnutrition Universal Screening Tool' ('MUST') for Adults. Redditch, Worchestershire, UK: Malnutrition Action Group. ISBN 978-1-899467-71-6

Basler HD, Bloem R, Casser HR et al. (2001) Ein strukturiertes Schmerzinterview für geriatrische Patienten. *Schmerz* 15:164-171

Beauchamp MK, Hao Q, Kuspinar A et al. (2023) A unified framework for the measurement of mobility in older persons. *Age Ageing* 52(Suppl 4): iv82-iv85. DOI: 10.1093/ageing/afad125

Bertoli S, Probst R, Jordan P (1996) Das Hörhandicap – Eine Ergänzung zum audiometrischen Hörverlust. *HNO* 44:376–384

Blazer D, Bachar JR, Hughes DC (1987) Major depression with melancholia. A comparison of middle-aged and elderly adults. *J Am Geriatr Soc* 35:927–932

Bowen M, Edgar DF, Hancock B et al. (2016) The prevalence of visual impairment in people with dementia (the ProVIDe study): a cross-sectional study of people aged 60–89 years with dementia and qualitative exploration of individual, carer and professional perspectives. *Health Serv Deliv Res* 4(21). DOI: 10.3310/hsdr04210

Bundesinstitut für Arzneimittel und Medizinprodukte (2026)
<https://klassifikationen.bfarm.de/ops/kode-suche/htmlops2026/block-8-55...8-60.htm> (Zugriff 29.01.2026)

BWKG-Arbeitshilfe. <https://www.bwkg.de/aufgaben-services/publikationen/identifikation-des-geriatrischen-patienten.html> (Zugriff 29.01.2026)

Cereda E (2012) Mini nutritional assessment. *Curr Opin Clin Nutr Metab Care* 15:29–41

Chalmers JM, King PL, Spencer AJ et al. (2005) The oral health assessment tool - Validity and reliability *Aust Dent J* 50(3):191-199. DOI: 10.1111/j.1834-7819.2005.tb00360.x

Cortes-Aguilar R, Malih N, Abbate M et al. (2024) Validity of nutrition screening tools for risk of malnutrition among hospitalized adult patients: A systematic review and meta-analysis. *Clin Nutr* 43:1094-1116. [https://www.clinicalnutritionjournal.com/article/S0261-5614\(24\)00082-7/fulltext](https://www.clinicalnutritionjournal.com/article/S0261-5614(24)00082-7/fulltext)

Courtin E & Knapp M (2017) Social isolation, loneliness and health in old age: a scoping review. *Health and Social Care in the Community* 25(3):799–812

Cudjoe TKM, Selvakumar S, Chung SE et al. (2022) Getting under the skin: Social isolation and biological markers in the National Health and Aging Trends Study. *J Am Geriatr Soc* 70(2):408-414. DOI: 10.1111/jgs.17518

Denkinger MD, Franke S, Rapp K et al. (2010) Accelerometer-based physical activity in a large

observational cohort - study protocol and design of the activity and function of the elderly in Ulm (ActiFE Ulm) study. BMC Geriatr 10 (50). <https://doi.org/10.1186/1471-2318-10-50>

Detsky AS, Mc Laughlin JR, Baker JP et al. (1987) What is Subjective Global Assessment of Nutritional Status? J Parent Enteral Nutr 11:8-13

Deutsche Schmerzgesellschaft e.V., Deutsche Gesellschaft für Geriatrie e.V., UVSD SchmerzLOS e.V. (Hrsg.) (2025): S3-Leitlinie Schmerzmanagement bei Geriatrischen Patient:innen in allen Versorgungssettings (GeriPAIN) <https://register.awmf.org/de/leitlinien/detail/145-005>

DGN e. V. & DGPPN e. V. (Hrsg.) S3-Leitlinie Demenzen, Version XX, 8.11.2023, <https://register.awmf.org/de/leitlinien/detail/038-013>

Dorresteijn TA, Zijlstra GR, K. Delbaere K et al. (2011) Evaluating an in-home multicomponent cognitive behavioural programme to manage concerns about falls and associated activity avoidance in frail community-dwelling older people: Design of a randomised control trial. BMC Health Services Research 11 (1):228 DOI 10.1186/1472-6963-11-228

Dräger D, Kreutz R, Wenzel A et al. (2021): Ältere Pflegebedürftige mit chronischen Schmerzen: Querschnittsstudie zur geschlechtsspezifischen Schmerzintensität und Versorgung in der großstädtischen Häuslichkeit. Schmerz 35 (5): 322–332. DOI: 10.1007/s00482-021-00538-5.

Drayer RA, Mulsant BH, Lenze EJ et al. (2005) Somatic symptoms of depression in elderly patients with medical comorbidities. International journal of geriatric psychiatry 20:973–982 DOI: 10.1002/gps.1389

van Dronkelaar C, Tieland M, Cederholm T et al. (2023) Malnutrition Screening Tools Are Not Sensitive Enough to Identify Older Hospital Patients With Malnutrition. Nutrients 15(24):5126. DOI:10.3390/nu15245126

Eid M, Diener E (Hg.) (2006) Handbook of multimethod measurement in psychology. 1. ed. Washington, DC: American Psychological Association

Feist B, Feindt B, Kasprick L et al (2022) Prüfung des Screening Tools GeriNOT. Identifikation geriatrischer Risikopotenziale bei Aufnahme ab 70-Jähriger in die Krankenhausversorgung. Z Gerontol Geriatr <https://doi.org/10.1007/s00391-022-02054-2>

Freedman A, Nicolle J (2020) Social isolation and loneliness: the new geriatric giants: Approach for primary care. Can Fam Physician 66(3):176-182

Freyenhagen R, Baron R, Gockel U, Tölle TR (2006) painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin 22:1911-1920

Gallo JJ, Bogner HR, Morales KH, Ford DE (2005) Patient ethnicity and the identification and active management of depression in late life. In: Archives of internal medicine 165:1962–1968 DOI: 10.1001/archinte.165.17.1962

Gerdes N, Funke UN, Schüwer U et al. (2012) „Selbständigkeits-Index für die Neurologische und Geriatrische Rehabilitation (SINGER)“ – Entwicklung und Validierung eines neuen Assessment-Instruments. Rehabilitation 51(05): 289-299

Guigoz Y, Vellas B, Garry PJ (1994) Mini nutritional assessment: A practical assessment tool for grading the nutritional state of elderly patients Facts, Research in Gerontology Suppl 2:15-59

Hadjistavropoulos T, Herr K, Prkachin KM et al. (2014) Pain assessment in elderly adults with dementia. The Lancet Neurology 13:1216-1227

- Hassel AJ, Rolko C, Koke U et al. (2008) A German version of the GOHAI. Community Dent Oral Epidemiol 36(1):34-42. DOI: 10.1111/j.1600-0528.2007.00351.x
- Hazan E, Frankenburg F, Brenkel M, Shulman K (2018) The test of time: a history of clock drawing. Int J Geriatr Psychiatry 33:e22–30. DOI: 10.1002/gpr.4731
- Hobert MA, Bernhard FP, Bettecken K et al. (2019) Validierung des Geriatrie-Checks in einer Kohorte von stationären neurologischen Patienten. Z Gerontol Geriatr 2:172-178 DOI: 10.1007/s00391-018-1441-5
- Hofmann W, Nikolaus T, Pientka L, Stuck AE (1995) Arbeitsgruppe Geriatisches Assessment (AGAST): Empfehlungen für den Einsatz von Assessment-Verfahren. Z Gerontol Geriatr 28:29–34
- Jahoda M (1958) Current concepts of positive mental health. In: New York, Basic Books
- Jeste DV, Blazer DG, First MB (2005): Aging-related diagnostic variations. Need for diagnostic criteria appropriate for elderly psychiatric patients. In: Biological psychiatry 58:265–271 DOI: 10.1016/j.biopsych.2005.02.004
- Jessen F, Spottke A, Deuschl G et al. (2017) S3-Leitlinie Demenzen. Springer Berlin Heidelberg. https://www.awmf.org/uploads/tx_szleitlinien/038-013I_S3-Demenzen-2016-07.pdf
- Junius-Walker U, Daether-Kracke N, Krause O (2016) It's Magic - einfaches geriatrisches Basisassessment für die Hausarztpraxis validiert. Z Allg Med 92:169-175
- Keith RA, Granger CV, Hamilton BB, Sherwin FS (1987) The functional independence measure: a new tool for rehabilitation. Adv Clin Rehabil. 1:6-18
- Kelly ME, Duff H, Kelly S et al. (2017) The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: a systematic review. Syst Rev 6, 259. <https://doi.org/10.1186/s13643-017-0632-2>
- Kiesswetter E, Sieber CC, Volkert D (2016) Ernährungssituation älterer Menschen im deutschsprachigen Raum. Aktuel Ernährungsmed 41:362–369
- Klotz AL, Zajac M, Ehret J et al. (2020) Development of a German version of the Oral Health Assessment Tool. Aging Clin Exp Res 32(1):165-172. doi: 10.1007/s40520-019-01158-x
- Kondrup J, Allison SP, Elia M et al. (2003) ESPEN guidelines for nutrition screening 2002. Clin Nutr 22:415-421
- Kruisbrink M, Crutzen R, Kempen GI, Zijlstra GA (2021) Assessing avoidance behavior due to concerns about falling: Psychometric properties of the FES-IAB in a sample of older adults of an online panel. Arch Gerontol Geriatr 97:104469 doi.org/10.1016/j.archger.2021.104469
- Krupp S, Kasper J, Hermes A et al. (2019) Das „Lübecker Modell Bewegungswelten“ – Ergebnisse der Effektevaluation. Bundesgesundheitsbl 62:274-281. <https://doi.org/10.1007/s00103-019-02881-6>
- Kudelka J, Ollenschläger M, Dodel R et al. (2024) Which Comprehensive Geriatric Assessment (CGA) instruments are currently used in Germany: a survey. BMC Geriatrics 24:347 <https://doi.org/10.1186/s12877-024-04913-6>
- Kuiper JS, Zuidersma M, Zuidema SU et al. (2016). Social relationships and cognitive decline: a systematic review and meta-analysis of longitudinal cohort studies, Int J Epidemiol 45(4):1169–1206. <https://doi.org/10.1093/ije/dyw089>
- Lachs MS, Feinstein AR, Cooney LM Jr. et al. (1990) A simple procedure for general screening for functional disability in elderly patients. Ann Intern Med 112:699–706

- Landesrahmenvereinbarung zur geriatrischen Versorgung in Schleswig-Holstein (2010) S. 5
https://www.schleswig-holstein.de/DE/Fachinhalte/K/krankenhaeuser/Downloads/krankenhaeuser_LRVgeriatrei2010.pdf?__blob=publicationFile&v=1
- Leitliniengruppe Hessen, DEGAM: S3-Leitlinie Multimedikation, Langfassung, AWMF-Registernummer: 053 – 043. 2. Auflage 2021
- Lezak M, Howieson DB, Bigler ED, Tranel D (2012) Neuropsychological Assessment. 5th. Edition. Oxford University Press
- Liebl ME, Elmer N, Schroeder I et al. (2016) Introduction of the Charité Mobility Index (CHARMI)– A novel clinical mobility assessment for acute care rehabilitation. PLoS One 11(12): e0169010. <https://doi.org/10.1371/journal.pone.0169010>
- Lichtner V, Dowding D, Esterhuizen P et al. (2014) Pain assessment for people with dementia: a systematic review of systematic reviews of pain assessment tools. BMC Geriatr 14:138
- Lin LC, Wu SC, Chen HS et al (2002) Prevalence of impaired swallowing in institutionalized older people in Taiwan. J Am Geriatr Soc 50:1118–1123
- Löhler J, Lehmann M, Segler V et al. (2019) Der Mini-Audio-Test (MAT) – eine Screeningmethode auf Schwerhörigkeit für Haus- und Fachärzte Laryngo-Rhino-Otol 98: 27–34. DOI: 10.1055/a-0805-5741
- Lukas A, Niederecker T, Gunther I et al. (2013) Self- and proxy report for the assessment of pain in patients with and without cognitive impairment: experiences gained in a geriatric hospital. Z Gerontol Geriatr 46:214-221
- Montero-Odasso M, van der Velde N, Martin FC et al.(2022) World guidelines for falls prevention and management for older adults: a global initiative *Age Ageing* 51(9) afac205, <https://doi.org/10.1093/ageing/afac205>
- Morello R, Jean A, Alix M et al. (2007) A scale to measure pain in non-verbally communicating older patients: the EPCA-2 Study of its psychometric properties. Pain 133
- Moser S, Luxenberger W, Freidl W (2017) Wahrnehmung von Hörproblemen bei älteren Menschen. HNO 65:671-679. DOI 10.1007/s00106-017-0334-z
- Nikolaus T, Specht-Leible N, Bach M et al. (1994) Soziale Aspekte bei Diagnostik und Therapie hochbetagter Patienten. Erste Erfahrungen mit einem neu entwickelten Fragebogen im Rahmen des geriatrischen Assessment. Z Gerontol 7:240–245
- Perez NB, Sierra PG, Taylor B, Fletcher J (2025) Selecting a depression measure for research: A critical examination of five common self-report scales. J Affect Disord 26:120542. DOI: 10.1016/j.jad.2025.120542
- Pirlich M, Schutz T, Norman K et al. (2006) The German hospital malnutrition study. Clin Nutr 25:563–572
- Quint S, Luckmann J, Wolf U et al. (2007) AMIKA: psychometric evaluation of a photo-based scale for the assessment of fear avoidance beliefs in elderly individuals. Schmerz 21:453-461
- Rasulo D, Christensen K, Tomassini C (2005) The Influence of Social Relations on Mortality in Later Life: A Study on Elderly Danish Twins. The Gerontologist 45(5):601–608. <https://doi.org/10.1093/geront/45.5.601>
- Ressler B, Ammer K (2010) Der ICIQ-SF als Messsystem zur Diagnose und Therapieverlauf der Harninkontinenz. Österr. Z. Phys. Med .Rehabil 20(1): 27-33
<http://www.uhlen.at/oezpmmr/data/pdf/200127.pdf>

- Rubenstein LZ, Josephson KR, Wieland GD et al. (1984) Effectiveness of a geriatric evaluation unit. A randomized clinical trial. *N Engl J Med* 311:1664-1670
- Runge M, Rehfeld G (2001) *Geriatrische Rehabilitation im Therapeutischen Team*. 2. Auflage. Thieme, Stuttgart
- Runge M, Wahl JM (Hrsg.) (1996) Lokomotion. In: *Ambulantes geriatrisches Assessment. Werkzeuge für die ambulante geriatrische Rehabilitation*. Steinkopff, Darmstadt, S. 23 - 26
- Scheffels JF, Kräling H, Kalbe E, Kessler J (2018) Konversionen von kognitiven Screenings. Mini-Mental-Status-Test vs. Montreal Cognitive Assessment vs. DemTect. *Nervenarzt* 89:1371-1377
- Schönle PW (1995) Der Frühreha-Barthelindex (FRB) - eine frührehabilitationsorientierte Erweiterung des Barthelindex. *Rehabilitation* 34:69-73
- Schütz T, Valentini L, Plaut M (2005) Screening auf Mangelernährung nach den ESPEN-Leitlinien 2002. *Aktuel Ernährungsmed* 30: 99-103 DOI: 10.1055/s-2004-834733
- Schwarzbach M, Luppä M, Forstmeier S et al. (2014) Social relations and depression in late life-a systematic review. *Int J Geriatr Psychiatry* 29(1):1-21. DOI: 10.1002/gps.3971
- Serra-Prat M, Palomera M, Gomez C et al. (2012) Oropharyngeal dysphagia as a risk factor for malnutrition and lower respiratory tract infection in independently living older persons: a population-based prospective study. *Age Ageing* 41:376–381
- Sirsch E unter Mitarbeit von Bartholomeyczik S, Halek M, Panke-Kochinke B (2020): Entscheidungsfindung zum Schmerzassessment bei Menschen mit Demenz im Krankenhaus. Vom Problem zum Konzept "See-Pain". ISBN 978-3-7799-3376-2z. Beltz Juventa in der Verlagsgruppe Beltz, Weinheim Basel
- Sirsch E, Gnass I, Laekeman M et al. (2018) Schmerzassessment bei älteren Menschen in der vollstationären Altenhilfe - Leitlinienreport der multiprofessionellen S3-Leitlinie, AWMF Registernummer 145-001. In: *Deutsche Schmerzgesellschaft, Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Witten*
- Spenciore B, Alves H, Charcat-Fishmann H (2017) Scoring systems for the Clock Drawing Test. A historical review. *Dement Neuropsychol* 11:6–14 DOI: 10.1590/1980-5764201dn11-01003
- Spix C, Blettner M (2012) *Dtsch Arztebl Int* 109:385-390 DOI: 10.3238/arztebl.2012.0385
- Šteinmiller J, Routasalo P, Suominen T (2015) Older people in the emergency department: a literature review. *Int J Older People Nurs* 10(4):284-305. DOI: 10.1111/opn.12090
- Stratton RJ, Hackston A, Longmore D et al. (2004) Malnutrition in hospital outpatients and inpatients: Prevalence, concurrent validity and ease of use of the 'malnutrition universal screening tool' ('MUST') for adults. *Br J Nutr* 92:799–808
- Thiem U, Greuel HW, Reingraber A et al. (2012) Positionspapier zur Identifizierung geriatrischer Patienten in Notaufnahmen in Deutschland. *Z Gerontol Geriat* 45:310-314
- Turley R, Cohen S (2009) Impact of voice and swallowing problems in the elderly. *Otolaryngol Head Neck Surg* 140:33–36
- van der Steen JT, Sampson EL, Van den Block L et al. (2015) Tools to Assess Pain or Lack of Comfort in Dementia: A Content Analysis. *J Pain Symptom Manage* 50:659-675 e653
- Ventry I, Weinstein BE (1982) The hearing handicap inventory for the elderly: a new tool. *Ear Hear* 3:128–134

- Volkert D, Beck AM, Cederholm T et al. (2019, epub 2018) ESPEN guideline on clinical nutrition and hydration in geriatrics. Clin Nutr 38:10-47 DOI: 10.1016/j.clnu.2018.05.024
- Volkert D, Bauer JM, Frühwald T et al. (2013) Leitlinie der Deutschen Gesellschaft für Ernährungsmedizin (DGEM) in Zusammenarbeit mit der GESKES, der AKE und der DGG. Klinische Ernährung in der Geriatrie. Aktuel Ernaehrungsmed 38:e1-e48
- Warden V, Hurley AC, Volicer L (2003) Development and psychometric evaluation of the Pain Assessment in Advanced Dementia (PAINAD) scale. J Am Med Dir Assoc 4:9-15
- Warren MW (1943) Care of Chronic Sick. Br Med J 2:822–823
- Wellner P (2012) Geriatrisches Screening in der Notaufnahme - Erfahrungen und erste Ergebnisse. Phys Med Rehab Kuror 22 - A80 DOI: 10.1055/s-0032-1322880
- WHO (2005) Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit. Deutsches Institut für Medizinische Dokumentation und Information, DIMDI (Hrsg)
https://web.archive.org/web/20150329082001/http://www.dimdi.de/dynamic/de/klassi/download_center/icf/endafassung/icf_endafassung-2005-10-01.pdf
- Wiedemann A für die Arbeitsgruppe Inkontinenz der Deutschen Gesellschaft für Geriatrie (2024) S2k-Leitlinie Harninkontinenz bei geriatrischen Patienten, Diagnostik und Therapie. AWMF-Registernummer 084-001 Langfassung
<https://register.awmf.org/de/leitlinien/detail/084-001> (Zugriff 14.11.2024)
- Wittich W, Phillips N, Nasreddine ZS, Chertkow H (2010). Sensitivity and Specificity of the Montreal Cognitive Assessment Modified for Individuals who are Visually Impaired. Journal of Visual Impairment & Blindness, 104(6):360-368. DOI: 10.1177/0145482X1010400606.
- Wolter DK (2011) Sucht im Alter – Altern und Sucht. Stuttgart. Kohlhammer
- Wolter DK (2016) Depressionen im höheren Lebensalter, Teil 1: Entstehung, klinische Symptome, Diagnose und Wechselwirkungen zwischen Depression und Demenz. Z Gerontol Geriatr. 49(4):335-48
- Wolter DK (2017) Schmerzen und Schmerzmittelabhängigkeit im Alter. Die gerontopsychiatrische Perspektive. Stuttgart. Kohlhammer
- World Health Organization (WHO) (2021) Social isolation and loneliness among older people: advocacy brief. ISBN 978-92-4-003074-9
- Yong B, Lin R, Xiao H (2020) Factors associated with nursing home adjustment in older adults: A systematic review, Int J of Nursing Studies 113(6):103790. DOI: 10.1016/j.ijnurstu.2020.103790
- Zietemann V, Kopczak A, Müller C et al. (2017) Validation of the Telephone Interview of Cognitive Status and Telephone Montreal Cognitive Assessment Against Detailed Cognitive Testing and Clinical Diagnosis of Mild Cognitive Impairment After Stroke. Stroke 48(11):2952-2957. DOI: 10.1161/STROKEAHA.117.017519.

4 Abkürzungen

AUC	→	area under the curve
CGA		Comprehensive Geriatric Assessment
CI	→	Confidence Interval
DSM	→	Diagnostic and Statistical Manual of Mental Disorders
ES	→	Effektstärke
F	→	Frau(en)
FEES	→	Flexible Endoscopic Evaluation of Swallowing
ICC	→	Intraclass Correlation Coefficient
J	→	Jahr(e)
m	→	Meter
MCI	→	Mild Cognitive Impairment
min	→	Minute
RCT	→	randomisierte kontrollierte Studie
ROC	→	Receiver Operating Characteristic Curve

Versionsnummer: 7.0

Erstveröffentlichung: 06/2019

Überarbeitung von: 14.02.2026

Nächste Überprüfung geplant: 14.02.2027

Die AWMF erfasst und publiziert die Leitlinien der Fachgesellschaften mit größtmöglicher Sorgfalt - dennoch kann die AWMF für die Richtigkeit des Inhalts keine Verantwortung übernehmen. **Insbesondere bei Dosierungsangaben sind stets die Angaben der Hersteller zu beachten!**

Autorisiert für elektronische Publikation: AWMF online