

Aus dem Institut für Allgemeinmedizin

Institut der Ludwig-Maximilians-Universität München



Kurze Instrumente zum Screenen von Suizidalität in der Hausarztpraxis:
Ein systematischer Review

Dissertation
zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

vorgelegt von

Milena Frank

aus

München

Jahr

2022

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

Erster Gutachter: Prof. Dr. med. Dipl.-Päd. Jochen Gensichen,
MPH

Zweiter Gutachter: Prof. Dr. med. Stefan Leucht

Dritter Gutachter: Prof. Dr. rer. nat. Ulrich Mansmann

Mitbetreuung durch den
promovierten Mitarbeiter: PD Dr. phil. Karoline Lukaschek

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 07.12.2022

A DANKSAGUNG

Ich bedanke mich herzlich bei meiner Betreuungskommission. Insbesondere bedanke ich mich bei meinem Doktorvater Professor Jochen Gensichen, der mir die Möglichkeit gegeben hat, diese Arbeit unter seiner Leitung durchzuführen und mir mit seiner Erfahrung und Ratschlägen zur Seite stand. PD Dr. Karoline Lukaschek danke ich für die hervorragende Betreuung und ihre ständige Diskussions- und Hilfsbereitschaft. Bei Professor Stefan Leucht bedanke ich mich für wertvolle Beiträge als Mitglied der Betreuungskommission.

Vielen Dank an Dr. Kathrin Halfter für die Prüfung meiner statistischen Berechnungen und die wissenschaftliche Zusammenarbeit im Rahmen der Publikation. Ich bedanke mich bei Professor Antonius Schneider für wertvolle Anregungen im Rahmen der Publikation. Bei Dr. Andreas Will bedanke ich mich für die Unterstützung bei der Literaturrecherche. Mein Dank gilt Dr. Linda Sanftenberg für allgemeine Tipps zur Erstellung der Dissertation und Daniela Lindemann für die Unterstützung bei technischen Fragen und Hilfe bei der Literaturbeschaffung. Dem Institut für Allgemeinmedizin danke ich für anregende Diskussionen im Rahmen des Work in Progress Seminars. Bei Sandra Nauerz bedanke ich mich für die schnelle und verlässliche Kommunikation.

Ich bedanke mich bei meiner Familie für die uneingeschränkte Unterstützung. Besonders danke ich meinen Eltern Maria Frank und Dr. Andreas Frank, die meinen bisherigen Lebensweg ermöglichten. Ich danke meinem Partner Dr. Oliver Jarosch für seine Liebe und seine Motivation.

B PUBLIKATION

1. Das Projekt wurde 2019 in einem Vortrag von Milena Frank auf dem 53. Kongress für Allgemeinmedizin und Familienmedizin der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) in Erlangen vorgestellt.

2. Das Projekt wurde in dem allgemeinmedizinischem Journal *MMW - Fortschritte der Medizin* veröffentlicht: Lukaschek K, Frank M, Halfter K, Schneider A, Gensichen J. (2021). Kurze Screener für Suizidalität in der Allgemeinmedizin. *MMW-Fortschritte der Medizin*, 163(6), 9-18.

Hierbei (Lukaschek K, Frank M, Halfter K, Schneider A, Gensichen J. (2021). Kurze Screener für Suizidalität in der Allgemeinmedizin. *MMW-Fortschritte der Medizin*, 163(6), 9-18.) handelt es sich um eine geteilte Erstautorschaft von PD Dr. phil. Karoline Lukaschek und Milena Frank. Milena Frank erstellte das Studienprotokoll, erarbeitete die Ein- und Ausschlusskriterien, entwickelte die Suchstrategie, führte die Literaturrecherche, Deduplikation, sowie Titel- und Abstractselektion durch, beschaffte die Volltexte, führte die Volltextselektion, sowie Datensammlung und -extraktion durch, führte die statistische Analyse, sowie die Qualitätsbewertung zum einen im Rahmen der Bewertung des Verzerrungsrisikos und Bedenken bezüglich der Anwendbarkeit, zum anderen die Bewertung der Gewissheit der Evidenz durch. Milena Frank erarbeitete die Datensynthese und erstellte das Manuskript, sowie jedwede Tabellen und Abbildungen der Publikation. In vorliegender Dissertation verwendete Tabellen, Abbildungen oder Textstellen wurden dementsprechend im fortlaufenden Text gekennzeichnet.

C ZUSAMMENFASSUNG

Hintergrund

Im Jahr 2020 verstarben in der Bundesrepublik Deutschland 9 206 Menschen an Suizid. Fast die Hälfte der Suizidenten besuchte im Monat vor dem Suizid einen Hausarzt. Dieser erkannte die Suizidalität seines Patienten oft nicht.

Daher stellt Screening einen ersten und wichtigen Schritt in der Diagnostik von Suizidalität dar. Ziel ist es, suizidale Patienten mit Hilfe kurzer sensitiver und spezifiver Fragebögen zu identifizieren.

Zielsetzung

Vorliegende Übersichtsarbeit gibt einen Überblick über vorhandene kurze Screening Instrumente für Suizidalität in der Allgemeinmedizin und Allgemeinbevölkerung und vergleicht diese hinsichtlich ihrer diagnostischen Genauigkeit.

Methode

Es wurde eine Literaturrecherche in den folgenden Datenbanken durchgeführt: MEDLINE, EMBASE, PSYINDEX, PsychINFO und Cochrane Library. Das Verzerrungsrisiko wurde anhand des QUADAS-2 Tool bewertet. Die Gewissheit der Evidenz wurde mit GRADE eingeschätzt. Der Einsatz der PRISMA Diagnostic Test Accuracy Checkliste diente der Qualitätssicherung. Ein Studienprotokoll wurde prospektiv auf PROSPERO veröffentlicht.

Ergebnisse

Insgesamt identifizierte die Suchstrategie N = 12 460 Artikel, von denen sieben Studien in die Übersichtsarbeit eingeschlossen werden konnten. Aufgrund der Heterogenität der Studien wurden diese in einer qualitativen Synthese zusammengefasst. In der Population der Allgemeinmedizin wurden drei und in der Allgemeinbevölkerung

vier Studien identifiziert, die kurze Screening Instrumente untersuchten. Unterschiedliche Ziel-Zustände wurden betrachtet (Suizidgedanken, Suizidplan, Suizidversuch, Suizidalität im Allgemeinen).

Die Prävalenz der Suizidalität wies in den Studien eine große Bandbreite von 1% bis 75% auf. Die Werte für Sensitivität bewegten sich zwischen 26% und 100%. Für die Spezifität ergaben sich Werte zwischen 64% und 99%. Der positive prädiktive Vorhersagewert bewegte sich zwischen 6% und 91%. Der negative prädiktive Vorhersagewert wies durchgehend Werte über 80% auf.

Aufgrund fehlender Untersuchung der diagnostischen Genauigkeit mussten acht kurze Screening Instrumente, die Suizidalität in der Allgemeinmedizin oder Allgemeinbevölkerung evaluierten, ausgeschlossen werden.

Conclusio

Obwohl ein Screening für Suizidalität in der Allgemeinmedizin bereits von unterschiedlichen Leitlinien empfohlen wird, konnten nur wenige Studien identifiziert werden, welche die diagnostische Genauigkeit eines Indextests untersuchten. Die identifizierten Indextests wurden in nur jeweils einer Studie untersucht und müssen ihre diagnostische Genauigkeit in weiteren Studien unter Beweis stellen. Eine Untersuchung der Screening Instrumente bei allgemeinmedizinischen Risikopatienten wäre sowohl aus statistischen Gründen als auch aufgrund der Empfehlung von Fachgesellschaften interessant. Das klinische Urteil des behandelnden Arztes stellt die endgültige Entscheidung dar.

D ABSTRACT

Background

In Germany 9 206 people died of suicide in 2020. Nearly half of the people who died by suicide visited a general practitioner a month before. General practitioners often did not recognize the suicidal behaviour of their patients.

Therefore, screening is a first and important step in the diagnosis of suicidal behaviour. The objective is to identify suicidal behaviour with the help of short sensitive and specific screening instruments.

Objective

The present systematic review provides an overview of existing short screening instruments for suicidal behaviour in primary care and the general population and compares them regarding their measures of diagnostic accuracy.

Methods

A literature search was carried out in the following databases: MEDLINE, EMBASE, PSYINDEX, PsychINFO and Cochrane Library. The risk of bias was examined using QUADAS-2 tool. The certainty of evidence was assessed with GRADE. The PRISMA Diagnostic Test Accuracy Checklist was used for quality management. A study protocol was published prospectively on PROSPERO.

Results

In total, the search strategy identified N = 12 460 articles, of which seven studies could be included in the review. Due to the heterogeneity of the studies, they were summarized in a qualitative synthesis. Three studies examined short screening instruments in primary care and four studies in the general population. Different target conditions were investigated (suicidal thoughts, suicide plan, suicide attempt and suicidal behaviour in general).

The prevalence of suicidal behaviour in the included studies ranged from 1% to 75%. The values for sensitivity ranged between 26% and 100%. The values for specificity ranged from 64% to 99%. The positive predictive value ranged between 6% and 91%. The negative predictive value was consistently above 80%.

Due to missing diagnostic accuracy, eight short screening instruments that evaluated suicidal behaviour in primary care or the general population had to be excluded.

Conclusion

Although screening for suicidal behaviour in primary care is already recommended by several guidelines, only a few studies could be identified that assessed diagnostic accuracy of an index test. The identified index tests were each examined in only one study and must prove their diagnostic accuracy in further studies. An investigation of the screening instruments in primary care patients at risk would be interesting for statistical reasons as well as for the recommendation of expert associations. The clinical judgement of the attending physician represents the final decision.

E INHALTSVERZEICHNIS

A	DANKSAGUNG	III
B	PUBLIKATION	IV
C	ZUSAMMENFASSUNG	V
D	ABSTRACT	VII
E	INHALTSVERZEICHNIS	IX
F	VERZEICHNISSE	XII
1	EINLEITUNG	1
1.1	ALLGEMEINES	1
1.2	BEGRIFFSDEFINITIONEN	1
1.3	EPIDEMIOLOGIE	2
1.4	RISIKO- UND SCHUTZFAKTOREN	3
1.5	PSYCHOLOGISCHE MODELLE DER SUIZIDALITÄT	5
1.6	SUIZIDPRÄVENTION	6
1.7	DIE BEDEUTUNG DER ALLGEMEINMEDIZIN IN DER SUIZIDPRÄVENTION	7
2	ZIELSETZUNG DER ARBEIT	10
3	MATERIAL UND METHODEN	11
3.1	DIE SYSTEMATISCHE ÜBERSICHTSARBEIT	11
3.2	STUDIEN ZUR DIAGNOSTISCHEN GENAUIGKEIT	12
3.3	VORRECHERCHER	12
3.4	QUALITÄTSSICHERUNG	13
3.5	DAS STUDIENPROTOKOLL	14
3.6	EIN- UND AUSSCHLUSSKRITERIEN	14
3.7	LITERATURERECHERCHER	15
3.7.1	<i>Informationsquellen</i>	15
3.7.2	<i>Suchstrategie</i>	16
3.8	STUDIENSELEKTION	18
3.8.1	<i>Deduplikation</i>	18
3.8.2	<i>Titel- und Abstractselektion</i>	18
3.8.3	<i>Volltextbeschaffung</i>	19

3.8.4	<i>Volltextselektion</i>	19
3.9	DATENSAMMLUNG UND DATENEXTRAKTION	20
3.10	STATISTISCHE METHODEN	20
3.11	QUALITÄTBEWERTUNG	21
3.11.1	<i>Verzerrungsrisiko und Bedenken bezüglich der Anwendbarkeit</i>	21
3.11.2	<i>Die Gewissheit der Evidenz (Certainty of evidence)</i>	22
3.12	DATENSYNTHESE.....	23
4	ERGEBNISSE	24
4.1	STUDIENSELEKTION	24
4.2	CHARAKTERISTIKA DER EINGESCHLOSSENEN STUDIEN	25
4.2.1	<i>Allgemeines</i>	25
4.2.2	<i>Teilnehmer der Studien</i>	29
4.2.3	<i>Indextest und Referenzstandard</i>	30
4.3	QUALITÄTBEWERTUNG	34
4.3.1	<i>Verzerrungsrisiko</i>	36
4.3.2	<i>Bedenken bezüglich der Anwendbarkeit</i>	37
4.3.3	<i>Grad der Übereinstimmung</i>	37
4.4	ERGEBNISSE DER EINZELNEN STUDIEN.....	37
4.5	GRUPPIERUNG DER ERGEBNISSE NACH DEM ZIEL-ZUSTAND	41
4.5.1	<i>Ziel-Zustand Suizidgedanken</i>	41
4.5.2	<i>Ziel-Zustand Suizidplan</i>	42
4.5.3	<i>Ziel-Zustand Suizidversuch</i>	42
4.5.4	<i>Ziel-Zustand Suizidalität</i>	43
4.6	SYNTHESE DER ERGEBNISSE	44
5	DISKUSSION	46
5.1	ALLGEMEINES.....	46
5.2	DISKUSSION DER METHODEN.....	46
5.2.1	<i>Methodische Stärken</i>	46
5.2.2	<i>Methodische Limitationen</i>	48
5.2.3	<i>Fazit</i>	50
5.3	DISKUSSION DER ERGEBNISSE.....	51
5.3.1	<i>Einordnung der Ergebnisse</i>	51
5.3.2	<i>Limitationen bezüglich der Ergebnisse</i>	59
5.3.3	<i>Fazit</i>	60

5.4	DIE BEDEUTUNG FÜR DIE ALLGEMEINMEDIZIN	67
5.4.1	<i>Herausstellungsmerkmale des Facharztes für Allgemeinmedizin</i>	67
5.4.2	<i>Die klinische Relevanz der Ergebnisse</i>	68
5.4.3	<i>Fachspezifische Limitationen von Screening für Suizidalität in der Allgemeinmedizin ...</i>	71
5.4.4	<i>Fazit und Forschungsaufgaben.....</i>	71
5.4.5	<i>Ausblick.....</i>	72
6	CONCLUSIO.....	74
6.1	SCHLUSSFOLGERUNG FÜR DIE FORSCHUNG.....	74
6.2	SCHLUSSFOLGERUNG FÜR DIE ALLGEMEINMEDIZINISCHE PRAXIS	74
G	ANHANG.....	XV
H	LITERATURVERZEICHNIS	XXI
I	AFFIDAVIT.....	XXXIII
J	CURRICULUM VITAE	XXXIV

F VERZEICHNISSE

Abkürzungsverzeichnis

C-SSRS	<i>Columbia - Suicide Severity Rating Scale</i>
DGS	<i>Deutsche Gesellschaft für Suizidprävention</i>
DIS	<i>Diagnostic Interview Schedule</i>
DTA	<i>Diagnostic Test Accuracy</i>
EQUATOR	<i>Enhancing the QUALity and Transperency Of health Research</i>
FN	<i>Falsch negativ</i>
FP	<i>Falsch positiv</i>
GHQ-28	<i>General Health Questionnaire - 28</i>
GRADE	<i>Grading of Recommendations Assessment, Development and Evaluation</i>
JG	<i>Prof. Dr. med. Dipl.-Päd. Jochen Gensichen, MPH</i>
K10	<i>Kessler Psychological Distress Scale</i>
KH	<i>Dr. rer. biol. hum., Dipl.-Biol. Kathrin Halfter</i>
KI	<i>Konfidenzintervall</i>
KL	<i>PD Dr. phil. Karoline Lukaschek</i>
MF	<i>Milena Frank</i>
NaSPro	<i>Nationales Suizidpräventionsprogramm für Deutschland</i>
NPV	<i>Negativer prädiktiver Vorhersagewert</i>
PHQ-9	<i>Patient Health Questionnaire - 9</i>
PPV	<i>Positiver prädiktiver Vorhersagewert</i>
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta Analyses</i>
RMTS-S	<i>Rapid Measurement Toolkit for Suicidality - Static Screener</i>
RN	<i>Richtig negativ</i>
RP	<i>Richtig positiv</i>
SAFE-T	<i>Suicide Assessment Five-step Evaluation and Triage</i>
SBQ-R	<i>Suicidal Behaviors Questionnaire - Revised</i>
SCD-I mood module	<i>Mood Module of the Structured Clinical Interview for DSM-IV Axis I Disorders</i>
SIDAS	<i>Suicidal Ideation Attributes Scale</i>
SIS-Q	<i>Suicidal Ideation Screening Questionnaire</i>
WHO	<i>Weltgesundheitsorganisation</i>

Abbildungsverzeichnis

<i>Abbildung 1: Durchführung einer systematischen Übersichtsarbeit in Anlehnung an Murad et al. (2014) (61)</i>	<i>11</i>
<i>Abbildung 2: PRISMA Flussdiagramm nach Vorlage von Moher et al. (2009) (103)</i>	<i>25</i>
<i>Abbildung 3: Geschlechterverteilung in den Studien der Allgemeinbevölkerung</i>	<i>29</i>
<i>Abbildung 4: Geschlechterverteilung in den Studien der Allgemeinmedizin.....</i>	<i>30</i>
<i>Abbildung 5: Alter der Studienteilnehmer</i>	<i>30</i>
<i>Abbildung 6: QUADAS-2 Domäne nach Vorlage der graphischen Darstellung der University of Bristol (120).....</i>	<i>35</i>
<i>Abbildung 7: Sensitivität und Spezifität in der Allgemeinbevölkerung dargestellt in einem Forest Plot</i>	<i>44</i>
<i>Abbildung 8: Sensitivität und Spezifität in der Allgemeinmedizin dargestellt in einem Forest Plot.....</i>	<i>45</i>
<i>Abbildung 9: Prozess eines Screenings für Suizidalität in der Allgemeinmedizin des Kaiser Permanente Washington Primary Care System (152).....</i>	<i>73</i>

Tabellenverzeichnis

<i>Tabelle 1: Risikofaktoren für Suizidalität in Anlehnung an O'Connor et al. (2014) (2)</i>	4
<i>Tabelle 2: Risikofaktoren für Suizidalität (10)</i>	4
<i>Tabelle 3: Schutzfaktoren für Suizidalität</i>	5
<i>Tabelle 4: Stress-Diathese-Modelle der Suizidalität</i>	6
<i>Tabelle 5: Ein- und Ausschlusskriterien der Studien</i>	15
<i>Tabelle 6: Datenbanken</i>	16
<i>Tabelle 7: Stichwortsuche</i>	17
<i>Tabelle 8: Thesaurussuche</i>	18
<i>Tabelle 9: Vierfeldertafel</i>	21
<i>Tabelle 10: Klassifizierung der Gewissheit der Evidenz. Übersetzt nach Tabelle „Information Certainty of evidence“ (100)</i>	23
<i>Tabelle 11: Charakteristika der eingeschlossenen Studien</i>	28
<i>Tabelle 12: QUADAS-2 Tool nach Vorlage der tabellarischen Darstellung der University of Bristol (119)</i>	35
<i>Tabelle 13: Die Ergebnisse der einzelnen Studien in der Allgemeinbevölkerung</i>	39
<i>Tabelle 14: Die Ergebnisse der einzelnen Studien in der Allgemeinmedizin</i>	40
<i>Tabelle 15: Erfassung der Risikomerkmale nach Empfehlung der S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression (40)</i>	54
<i>Tabelle 16: Überblick über die Einordnung der Ergebnisse</i>	58
<i>Tabelle 17: Studien ausgeschlossen aufgrund fehlender Maße zur diagnostischen Genauigkeit</i>	66

1 EINLEITUNG

1.1 ALLGEMEINES

Die Weltgesundheitsorganisation (WHO) fordert eine Entkriminalisierung und Entstigmatisierung von Suizidalität sowie einen höheren gesundheitspolitischen Stellenwert des Themas (1).

Die Suizidforschung hat sich in den letzten Jahrzehnten stark weiterentwickelt. Dabei wurden die Epidemiologie sowie Risiko- und Schutzfaktoren untersucht und kulturelle Unterschiede identifiziert (1). Ferner wurden unterschiedliche psychologische Modelle von Suizidalität entwickelt (2), welche die Komplexität des Forschungsgegenstandes widerspiegeln. Darüber hinaus wurden im Rahmen der Suizidprävention verschiedene Strategien evaluiert (1). Dabei nimmt Screening von Patienten einen wichtigen Stellenwert ein (3), und wird in diesem systematischen Review genauer behandelt: Screening stellt in der Diagnostik von Suizidalität einen ersten und unverzichtbaren Schritt dar. Ziel ist es, suizidale Patienten mit Hilfe kurzer sensitiver und spezifiver Fragebögen zu identifizieren (4).

Diese Übersichtsarbeit gibt einen Überblick über vorhandene kurze Screening Instrumente für Suizidalität in der Allgemeinmedizin und Allgemeinbevölkerung und vergleicht diese anhand ihrer diagnostischen Genauigkeit.

1.2 BEGRIFFSDEFINITIONEN

Suizid: „Tod aufgrund eines intentionalen, selbstschädigenden Verhaltens, das mit einem gewissen Maß an Absicht zu sterben assoziiert war.“ (5)

Suizidalität: „Suizidalität ist die Summe aller Denk- und Verhaltensweisen von Menschen oder Gruppen von Menschen, die in Gedanken, durch aktives Handeln, Handeln lassen oder passives Unterlassen den eigenen Tod anstreben bzw. als möglichen Ausgang einer Handlung in Kauf nehmen.“ (6)

Allgemeinmedizin: „Der Arbeitsbereich der Allgemeinmedizin beinhaltet die Grundversorgung aller Patienten mit körperlichen und seelischen Gesundheitsstörungen in der Notfall-, Akut- und Langzeitversorgung sowie wesentliche Bereiche der Prävention und Rehabilitation. Allgemeinärztinnen und Allgemeinärzte sind darauf spezialisiert, als erste ärztliche Ansprechpartner bei allen Gesundheitsproblemen zu helfen.“ (7)

1.3 EPIDEMIOLOGIE

Weltweit wird die Anzahl der Sterbefälle durch Suizid auf über 800 000 Menschen pro Jahr geschätzt. Das bedeutet, dass alle 40 Sekunden ein Mensch durch Suizid verstirbt. (1) Im Jahr 2020 verstarben in der Bundesrepublik Deutschland 9 206 Menschen an Suizid (8).

Die höchsten Suizidraten wurden zum einen bei Jugendlichen und jungen Erwachsenen (9), zum anderen bei Menschen höheren Alters festgestellt (9, 10). Bei Jugendlichen und jungen Erwachsenen stellte Suizid die zweithäufigste Todesursache dar (1).

Trotz dieser hohen Sterbeziffer vermutet *Hawton et al. (2009)* (10) eine weltweite Unterschätzung der Suizidrate. Dies führt der Autor auf unterschiedliche Klassifikationssysteme der Todesursachen zurück (unterschiedliche Definitionen, unterschiedliche Berufsgruppen), ferner auf die Kriminalisierung des Suizides in Ländern, wie z.B. Indien oder eine fehlende Datenerhebung, vor allem auf dem afrikanischen Kontinent. (10)

Kulturelle Unterschiede wurden bezüglich der Suizidraten berichtet (10). Eine Studie konnte z.B. eine höhere Suizidrate von Menschen kaukasischer Herkunft als von Menschen afrikanischer Herkunft feststellen (11). Weiterhin wurde ein saisonaler Anstieg

der Suizidrate im Frühjahr beschrieben (12). Auch das Geschlecht scheint einen Einfluss auf Suizid zu haben: Zum einen waren mehr als zwei Drittel der Suizidopfer Männer (13). Zum anderen wählten Männer häufiger sog. *harte* Suizidmethoden (14), die einen großen körperlichen Schaden anrichten und bei denen eine Rettung höchst unwahrscheinlich ist (15) (z.B. Erhängen, Erschießen (14), Bahnsuizid (15)). Frauen dagegen begingen öfter einen *weichen* Suizid (z.B. Tabletten) (14), bei dem eine rechtzeitige Rettung möglich sein kann. In den westlichen Industrienationen wurde die Hälfte der Suizide durch Erhängen ausgeführt (1).

Die Anzahl der Suizidversuche übersteigt die der Suizide stark (9). Die WHO berichtete über ein Verhältnis von 20:1 (1). Bei ungefähr einem Drittel der Menschen mit Suizidgedanken folgt ein Suizidversuch. In zwei Drittel der Suizidversuche geschieht dies bereits im ersten Jahr nach Einsetzen der Suizidgedanken (16).

In einer großen epidemiologischen Studie wurde eine Lebenszeitprävalenz von Suizidgedanken von 8% - 9% und eine Lebenszeitprävalenz von Suizidversuchen von 1% - 3% berichtet (16, 17). *Lukaschek et al. (2015)* berichtete eine Prävalenz von Suizidgedanken im vorangegangenen Monat von 9% (18). In der Allgemeinmedizin zeigte ein Review Prävalenzen von 1% - 11% (19).

1.4 RISIKO- UND SCHUTZFAKTOREN

Ärzte werden im Berufsalltag mit Patienten konfrontiert, bei denen sie Suizidalität einschätzen müssen (20). Hierfür ist die Kenntnis bedeutender Risikofaktoren wichtig. Bei Suizidalität handelt es sich um ein komplexes Phänomen (2), welches nicht leicht prognostizierbar ist und durch unterschiedliche Risiko- und Schutzfaktoren beeinflusst wird. *Tabelle 1* zeigt die vier Gruppen der Risikofaktoren nach *O'Connor et al. (2014)* (2):

Tabelle 1: Risikofaktoren für Suizidalität in Anlehnung an O'Connor et al. (2014) (2)

Individuell	Sozial
Hoffnungslosigkeit Impulsivität Perfektionismus	Suizid in der Familienanamnese Berichterstattung der Medien Fehlende soziale Unterstützung Soziale Isolation
Kognitiv	Negative Lebensereignisse
Kognitive Starrheit Gedankenkreisen Gefühl eine Last zu sein Schwierigkeiten Probleme zu lösen	Sexueller und körperlicher Missbrauch Physische Erkrankungen

Zentrale Risikofaktoren stellen insbesondere ein vorangegangener Suizidversuch sowie Suizidgedanken dar (1, 21-23). Rund 90% der Suizidenten weisen eine psychiatrische Erkrankung auf (24), wobei affektive Störungen und Schizophrenie einen großen Anteil in Anspruch nehmen (9, 24). Treffen unterschiedliche Risikofaktoren zusammen, erhöht das die Wahrscheinlichkeit von Suizidalität für die betroffene Person (1, 10).

Eine weitere geläufige Einteilung der Risikofaktoren stellt die Unterteilung in proximale und distale Faktoren dar (9, 10), siehe *Tabelle 2*. Dabei beschreiben die distalen Faktoren die Prädisposition eines Menschen und die proximalen Faktoren sogenannte auslösende Faktoren (9).

Tabelle 2: Risikofaktoren für Suizidalität (10)

Distal	Proximal
Genetik Neurobiologische Dysfunktion Perinatale Wachstumsverzögerung Frühe traumatische Lebensereignisse	Psychische Erkrankungen Psychosoziale Krisen Verfügbarkeit von Mitteln und Methoden zur Selbsttötung Physische Erkrankungen

Des Weiteren werden folgende Schutzfaktoren in der Wissenschaft beschrieben (*Tabelle 3*), jedoch besteht diesbezüglich weiterer Forschungsbedarf (2, 5).

Tabelle 3: Schutzfaktoren für Suizidalität

Individuell	Sozial	Religion
Optimismus (25-27) Resilienz (28)	Zugehörigkeit Verantwortung Unterstützung (5, 29)	(5, 29)

1.5 PSYCHOLOGISCHE MODELLE DER SUIZIDALITÄT

Tabelle 4 gibt einen Überblick über gängige psychologische Modelle der Suizidalität. Die meisten dieser Modelle basieren auf dem Diathese-Stress-Modell. Hierbei besteht die Annahme, dass sowohl eine Neigung (im Sinne einer Prädisposition) als auch Stress zu Suizidalität führen. Ein Modell, das verschiedene Handlungsphasen der Suizidalität (von passivem Todeswunsch hin zu aktiven Suizidgedanken und -handlungen) beschreibt und auch Lebensumstände und Umfeld der Person einbezieht, ist das integrativ motivationale-volitionale Modell suizidalen Verhaltens (30). Für den Allgemeinmediziner können diese verschiedenen Phasen und Aspekte in der Therapie des Patienten und der Suizidprävention wichtig sein. *Tabelle 4* wurde in Anlehnung an *O'Connor et al. (2014)* und *Teismann et al. (2016)* erstellt (2, 5).

Tabelle 4: Stress-Diathese-Modelle der Suizidalität

Autor, Jahr	Psychologisches Modell	Theorie
Baumeister, 1990 (31)	Escape-Theorie	Suizid dient der Flucht vor einer schmerzhaften Selbstwahrnehmung.
Williams, 2001 (32)	Cry of Pain-Modell	Erlebnisse, die als Scheitern wahrgenommen werden (defeat) und die als nicht entzerrbar scheinen (entrapement), führen zu Suizidalität.
Joiner, 2005 (33)	Interpersonale Theorie suizidalen Verhaltens	Suizidalität entsteht durch Interaktion folgender drei Faktoren: Einsamkeit (thwarted belongingness), das Gefühl eine Last zu sein (perceived burdensomeness) und einer gewissen Furchtlosigkeit vor dem Tod (capability for suicide).
Rudd, 2006 (34)	Fluid Vulnerability Theory of Suicide	Ein suizidaler Modus kann abhängig von bestehenden Vulnerabilitäten durch interne und externe Stressoren aktiviert werden.
Wenzel und Beck, 2008 (35)	Kognitive Theorie suizidalen Verhaltens	Unaushaltbarkeitsüberzeugungen münden getriggert durch Stress in akute Hoffnungslosigkeit. Selektive Kognitionen engen das Blickfeld auf Suizid als einzige Handlungsoption ein.
O'Connor, 2011 (30)	Integrativ motivationales-volitionales Modell suizidalen Verhaltens	Das Modell beschreibt die prämotivationalen, motivationalen (Suizidgedanken) und volitionalen (Suizidale Handlungen) Phasen der Suizidalität.

1.6 SUIZIDPRÄVENTION

Grundvoraussetzung einer erfolgreichen Suizidprävention sind Früherkennung und Intervention. Unter Berücksichtigung der multifaktoriellen Ätiologie und Genese von Suizidalität sind breitgefächerte Präventionsansätze erforderlich (1, 36).

Laut Expertenkonsens gibt es folgende fünf Hauptmaßnahmen: 1. Screening von Risikopatienten (37, 38). 2. Sensibilisierung und Aufklärungskampagnen sowie Schulungsprogramme (39) der Hausärzte, der Allgemeinbevölkerung und von Gatekeepern. 3. Die Behandlung von psychischen Erkrankungen durch eine vermehrte Nachsorge suizidgefährdeter Patienten (40), Psychotherapie und geeignete Medikation (37,

38). 4. Ein Erschweren des Zugangs zu tödlichen Mitteln und Methoden (36). 5. Leitlinien für die Berichterstattung von Selbsttötung in den Medien (41) zur Verringerung des Werther-Effekts (Nachahmung von Suiziden durch Berichterstattung in den Medien). (3)

In 28 Ländern gibt es bereits eine nationale Präventionsstrategie (1). Unter dem Nationalen Suizidpräventionsprogramm für Deutschland (NaSPro) (42), einer Initiative der Deutschen Gesellschaft für Suizidprävention (DGS) (43), haben sich über 90 Institutionen aus Politik, Kirche, Medien, Betroffenenverbände, Gesundheitswesen und Wissenschaft zu einem Netzwerk zusammengeschlossen. Die NaSPro stellt unter anderem eine Checkliste für die Berichterstattung von Suiziden zur Verfügung und ist Kooperationspartner des Projekts *Suizidprävention Deutschland – Aktueller Stand und Perspektiven*, gefördert durch das Bundesministerium für Gesundheit (44). Zudem ist sie an der Organisation des Welttages für Suizidprävention beteiligt (42).

Nur ein geringer Anteil der Menschen begeht einen Suizid ohne Warnsignale (1). Daher ist eine Prävention möglich und nötig.

1.7 DIE BEDEUTUNG DER ALLGEMEINMEDIZIN IN DER SUIZIDPRÄVENTION

Der Hausarzt kann in der Suizidprävention eine zentrale Schlüsselrolle einnehmen (1). Die Strategien zur Suizidprävention in der Allgemeinmedizin zielen auf Screening von Suizidalität (37), Trainings- und Sensibilisierungsprogramme für Hausärzte sowie die Behandlung von psychischen Erkrankungen ab (3).

Bis zu 80% der Suizidenten kontaktierten ihren Hausarzt im Jahr vor dem Suizid, ungefähr die Hälfte im Monat des Suizids (45, 46). Die Häufigkeit der Hausarztbesuche in der Zeit vor dem Suizid lag deutlich höher als bei Kontrollgruppen in der Allgemeinbevölkerung (47). Allerdings wurde die Suizidalität oft nicht erkannt (19, 47, 48). Nur eine Minderheit der Patienten besuchte spezialisierte psychiatrische oder psychotherapeutische Einrichtungen (45, 48).

Studien über Suizidalität in der Primärversorgung zeigten, dass etwa 1% bis 11% der Patienten Suizidgedanken hatten (19), aber nur ein geringer Anteil diese von sich aus ansprach (9, 49, 50). Dies deutet darauf hin, dass Patienten, die sich Hilfe gewünscht hätten, ihre Suizidgedanken nicht äußerten.

Die hohen Suizidraten einerseits und die mangelnde Diagnostik von Suizidalität andererseits werden in der öffentlichen Gesundheit kontrovers diskutiert. *Lukaschek et al. (2012)* konnte zeigen, dass Suizid mit einem großen Verlust an Lebensjahren einhergeht (51). Zu den wichtigsten Risikofaktoren für suizidales Verhalten, die bereits in der Allgemeinmedizin spezifisch untersucht wurden, gehören Angststörungen und Depressionen (52, 53).

Da die psychosomatische Grundversorgung einen essentiellen Bestandteil der Weiterbildung eines Allgemeinmediziners darstellt (54), muss ihm die Kompetenz, Suizidalität zu diagnostizieren, zugesprochen werden. Die Weiterbildung beinhaltet Differentialdiagnostik psychischer Krankheitsbilder, grundlegende Therapien und die Weitervermittlung derjenigen Patienten, die von einer psychotherapeutischen oder psychiatrischen Anbindung profitieren (54). Des Weiteren wird im *Kompetenzbasierten Curriculum Allgemeinmedizin*, welches dem Arzt in Weiterbildung als Leitfaden dient, explizit beschrieben, dass dieser psychische Notfälle wie einen drohenden Suizid diagnostizieren, sowie eine weitere Therapie einleiten können sollte (55).

Daher sollte ein aktives Screenen nach suizidalen Gedanken und Verhaltensweisen ein integraler Bestandteil der Allgemeinmedizin sein, insbesondere bei Risikopatienten (37, 38). Kurze Screening Instrumente könnten zeit- und kostenschonend in der Allgemeinmedizin eingesetzt werden (56). Folgende Vorteile von Messinstrumenten zur Einschätzung von Suizidalität werden von *Forkmann et al. (2015)* betont (4):

- „Screeninginstrumente sind ökonomisch einsetzbar und können den Weg zu einer ausführlicheren Risikoabschätzung bahnen.“
- „Selbstbeurteilungsinstrumente können die Risikoabschätzung ergänzen, insbesondere da es manchem Betroffenen leichter fallen mag, suizidale Erlebensweisen zunächst auf Papier zu offenbaren und nicht im unmittelbaren Gespräch.“
- „Interviewverfahren und Fremdbeurteilungsinstrumente können sicherstellen,

dass – insbesondere bei therapeutenseitigen Ängsten im Umgang mit Suizidalität – keine zentralen Informationen übersehen werden.“

- „Messinstrumente können für die Verlaufskontrolle und die Dokumentation genutzt werden und bieten damit auch unter Haftungsgesichtspunkten einen Mehrwert.“
- „Die Verwendung reliabler und valider Messinstrumente bildet die Basis jeglicher Forschungsaktivität.“

2 ZIELSETZUNG DER ARBEIT

Ziel dieses systematischen Reviews ist es, einen Überblick über den aktuellen Forschungsstand der Suizidpräventionsstrategie Screening mittels Screening Instrument auf dem Gebiet der Allgemeinmedizin zu geben. Hierfür wurden kurze Screening Instrumente für Suizidalität in der Allgemeinmedizin oder Allgemeinbevölkerung identifiziert und hinsichtlich ihrer diagnostischen Genauigkeit verglichen. Sollten die Forschungsergebnisse es zulassen, könnte eine Empfehlung eines Instruments für die Allgemeinmedizin gegeben werden. Weiterer Forschungsbedarf sollte aufgezeigt werden.

3 MATERIAL UND METHODEN

3.1 DIE SYSTEMATISCHE ÜBERSICHTSARBEIT

Die Anzahl an wissenschaftlichen Publikationen nimmt stetig zu (57). Ärzte haben oft nicht die Zeit, jede einzelne Studie zu lesen, zu vergleichen und zu bewerten und so eine evidenzbasierte Entscheidung zu treffen (58). Eine systematische Übersichtsarbeit soll den aktuellen Forschungsstand zu einer bestimmten Fragestellung klären (57). Für ein ausgewähltes Thema werden Studien gesammelt und kritisch bewertet. Durch die strukturierte Literaturrecherche soll gewährleistet werden, dass alle Studien zum gewählten Thema in die Auswertung einfließen (57).

Folgende *Abbildung 1* stellt den Prozess einer systematischen Übersichtsarbeit dar.

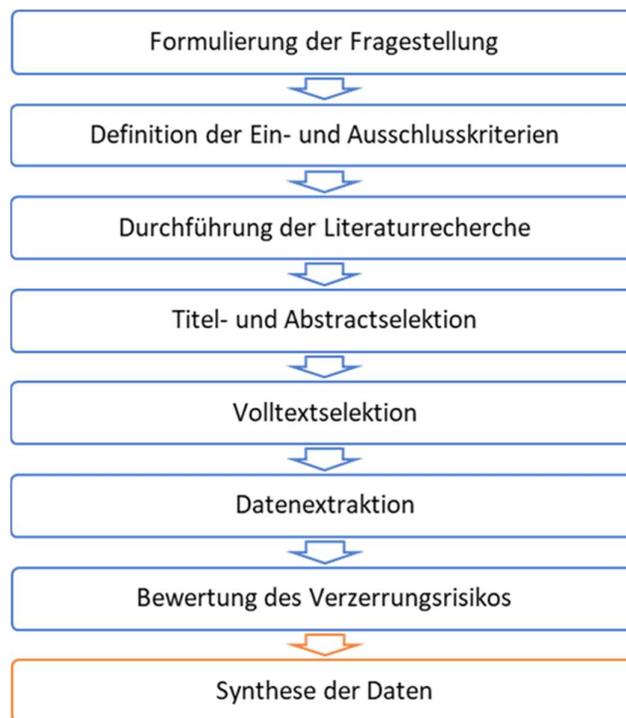


Abbildung 1: Durchführung einer systematischen Übersichtsarbeit in Anlehnung an Murad et al. (2014) (58)

3.2 STUDIEN ZUR DIAGNOSTISCHEN GENAUIGKEIT

In dieser systematischen Übersichtsarbeit wird eine Fragestellung zu diagnostischen Tests behandelt. Damit verbunden sind spezifische Besonderheiten im Gegensatz zu den geläufigeren Fragestellungen zu Interventionen.

Studien zur diagnostischen Genauigkeit untersuchen Tests hinsichtlich der Fähigkeit, Studienteilnehmer mit einem bestimmten Ziel-Zustand und Studienteilnehmer ohne diesen Ziel-Zustand voneinander zu differenzieren (59, 60).

Dabei werden die Ergebnisse eines Indextests (der Test, der in der Studie untersucht wird) mit einem Referenzstandard (vormals bekannt als Goldstandard) verglichen. Im Idealfall handelt es sich um Querschnittsstudien, in denen dieselben Teilnehmer der Studie beide Tests erhalten. In der Literaturrecherche sollen die Wortfelder *Krankheit*, *Test* und *Diagnostische Genauigkeit* mit dem Booleschen Operator *Und* verknüpft werden. Die graphische Synthese beinhaltet einen Plot für Sensitivität und Spezifität einschließlich Konfidenzintervall. (61)

Studien zur diagnostischen Genauigkeit stellen einen Standardstudientyp dar (59, 60). Ein Leitfaden für systematische Übersichtsarbeiten von Studien zur diagnostischen Genauigkeit der Cochrane Collaboration befindet sich bereits in Entwicklung: Das *Cochrane handbook for systematic reviews of diagnostic test accuracy*. Aktuell steht bereits ein Entwurf mehrerer Kapitel zur Verfügung. Eine erste Buchausgabe soll 2022 veröffentlicht werden. (62)

3.3 VORRECHERCHE

In einer ausführlichen Vorrecherche wurde das Themenfeld *Instrumente zum Screenen von Suizidalität in der Allgemeinmedizin* untersucht. Unterschiedliche Übersichtsarbeiten wurden identifiziert sowie als Orientierung verwendet. Einen Überblick über verschiedene Screener gaben das Standardwerk *A review of suicide assessment measures for intervention research with adults and older adults* von *Brown et al. (2001)* (63) als auch das Buch *Diagnostik von Suizidalität* (4). Identifizierte Reviews

(63-74) beinhalteten unterschiedliche Settings. Exemplarisch zu nennen waren hier psychiatrisches Setting, Allgemeinbevölkerung, Notaufnahme, ambulantes Setting und Allgemeinmedizin. Die Übersichtsarbeiten gaben einen Überblick über vorhandene Screening Instrumente oder bewerteten diese zusätzlich anhand unterschiedlicher Kriterien: Exemplarisch zu nennen sind psychometrische Eigenschaften, diagnostische Genauigkeit und freie Verfügbarkeit. Unterschiedliche Outcomes beinhalteten Suizidgedanken, Suizidplan, Suizidversuch, Suizidrisiko, Suizidalität im Allgemeinen und Suizid.

Die U.S. Preventive Services Task Force veröffentlichte bereits im Jahr 2013 eine Übersichtsarbeit mit der Untersuchung des Forschungsgegenstandes *Screening Instrumente für Suizidalität in der Allgemeinmedizin* (74).

Der vorliegende Review grenzte sich jedoch in folgenden Punkten zum Review der U.S. Preventive Services Task Force ab: Es handelte sich um den ersten Review dieser Art, der deutsche Publikationen berücksichtigte. Zudem fand eine Fokussierung auf kurze Screener mit höchstens zwölf Items statt (75). Eingeschlossen wurden ausschließlich Studien deren Studienteilnehmer eine große Altersspanne aufwiesen. Zusätzlich wurde die Population Allgemeinbevölkerung eingebunden, da aufgrund der Vorrecherche nur eine geringe Anzahl an Studien in der Allgemeinmedizin zu erwarten war und beide Populationen eine niedrige Prävalenz bezüglich Suizidalität aufwiesen (16, 19).

Das internationale prospektive Register systematischer Übersichtsarbeiten PROSPERO wurde auf bereits angemeldete Studien zu diesem Thema geprüft (76).

3.4 QUALITÄTSSICHERUNG

Das Enhancing the QUALity and Transparency Of health Research (EQUATOR) Netzwerk stellt Berichterstattungsleitlinien für wichtige Studienarten im Gesundheitswesen zur Verfügung (77) und dient somit der evidenzbasierten Wissenschaft. Dies beinhaltet einen Leitfaden für das Berichten von systematischen Übersichtsarbeiten, das Pre-

ferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) Statement (78). In dieser Übersichtsarbeit wurde eine Erweiterung, das PRISMA Statement für Diagnostic Test Accuracy (DTA) Studies, für Studien zur diagnostischen Genauigkeit angewendet (79).

Die Einhaltung der Leitlinie wurde bereits im Studienprotokoll verankert und während des Review-Prozesses stets berücksichtigt. Somit konnten eine Qualitätssicherung sowie Transparenz der Berichterstattung garantiert werden. Anhand der PRISMA-DTA Checkliste wurden Kernaspekte der Übersichtsarbeit erfasst. Die 27 Items umfassende Checkliste befindet sich mit den jeweiligen Seitenangaben im Anhang G (79).

3.5 DAS STUDIENPROTOKOLL

Das Studienprotokoll stellt einen bedeutenden Meilenstein in der Erstellung einer systematischen Übersichtsarbeit dar (57, 80). Es dient der Schärfung des Themas und leitet durch die Durchführung der wissenschaftlichen Arbeit (80).

Für die Veröffentlichung von Studienprotokollen systematischer Übersichtsarbeiten steht das internationale prospektive Register systematischer Reviews PROSPERO zur Verfügung. Da die systematischen Übersichtsarbeiten zu Beginn der Forschungsphase registriert werden müssen, kann eine Verzerrung der Berichterstattung verringert werden (Reporting Bias). Außerdem wird so die Wahrscheinlichkeit minimiert, dass dasselbe Thema mehrmals in einem Forschungsprojekt bearbeitet wird. (76)

Ein ausführliches Studienprotokoll wurde auf PROSPERO veröffentlicht (ID: CRD42019122173) (81).

3.6 EIN- UND AUSSCHLUSSKRITERIEN

Die Ein- und Ausschlusskriterien wurden anhand des PICO(S) Schema, einem Modell der evidenzbasierten Medizin erarbeitet (82) und im Studienprotokoll festgehalten.

Das Akronym steht für *P Participants, I Intervention, C Comparison, O Outcome* und *S Study*.

Um eingeschlossen zu werden mussten die Studienteilnehmer 14 Jahre oder älter sein, sowie der Allgemeinbevölkerung entstammen, oder Patienten der Allgemeinmedizin sein. Es wurden nur diese Studien eingeschlossen, die eine breitere Altersspanne der Studienteilnehmer aufwiesen. Der Indextest sollte Suizidalität messen und nicht mehr als zwölf Items aufweisen (75). Der Referenzstandard sollte Suizidalität messen und als Vergleich dienen. Die Tests konnten entweder als Selbstbeurteilungsfragebogen oder als Fremdbeurteilungsfragebogen fungieren. Die unterschiedlichen Screener wurden hinsichtlich ihrer diagnostischen Genauigkeit verglichen. (83) Weitere Details sind in *Tabelle 5* dargestellt.

Tabelle 5: Ein- und Ausschlusskriterien der Studien

	Einschlusskriterien	Ausschlusskriterien
Studienteilnehmer	14 Jahre und älter Allgemeinmedizin Allgemeinbevölkerung Mehrheit kaukasisch	Unter 14 Jahren Cognitives Impairment oder Demenz
Indextest	Screener für Suizidalität Bis zu zwölf Items	Mehr als zwölf Items
Vergleich	Maße der diagnostischen Genauigkeit	
Studiendesign	Diagnostische Studien Randomisierte kontrollierte Studien	Fall-Kontroll-Studien
Sprache	Englisch Deutsch	Nicht-Englisch Nicht-Deutsch
Status der Veröffentlichung	Peer-reviewed Artikel Bücher Dissertationen	Unveröffentlichte Studien

3.7 LITERATURERECHERCHE

3.7.1 INFORMATIONSMQUELLEN

Den Großteil an Informationsquellen stellten elektronische Datenbanken aus dem medizinischen und psychologischen Sektor dar. Die elektronische Literaturrecherche

wurde von einem erfahrenen Bibliothekar begleitet und in den folgenden Datenbanken durchgeführt: MEDLINE (84), EMBASE (84), Cochrane Library (85), PSYINDEX (86) und PsychINFO (86). Die durch die Universitätsbibliothek der Ludwig-Maximilians-Universität lizenzierten Zeitlimits wurden vollständig angewendet. Die endgültige Suchanfrage wurde am 25. Juli 2021 gestellt. Für weitere Informationen siehe *Tabelle 6*. Zusätzlich wurde eine manuelle Recherche durchgeführt. Dabei wurden Quellenangaben der eingeschlossenen Studien sowie relevante Übersichtsarbeiten und Bücher berücksichtigt. (83)

Tabelle 6: Datenbanken

Datenbank	Plattform	Zeitlimit
MEDLINE	Ovid	
Stichwortsuche		Ab 1946
Thesaurussuche		Ab 1946
EMBASE	Ovid	
Stichwortsuche		Ab 1974
Thesaurussuche		Ab 1974
Cochrane Library	Wiley Online Library	
Stichwortsuche		Kein Zeitlimit
Thesaurussuche		Kein Zeitlimit
PsychINFO	EBSCO	
Stichwortsuche		Ab 1938
Thesaurussuche		Ab 1977
PSYINDEX	EBSCO	
Stichwortsuche		Ab 1977
Thesaurussuche		Ab 1985

3.7.2 SUCHSTRATEGIE

Die Suchformel wurden in Anlehnung an *Glasziou et al. (2001)* (61) entwickelt. Es wurde eine Stichwortsuche für Titel und Abstract und eine Thesaurussuche durchgeführt. Die Wortfelder *Krankheit*, *Test* und *Diagnostische Genauigkeit* wurden mit dem Booleschen Operator *Und* verknüpft (61). Die Verknüpfung der Begriffe innerhalb der jeweiligen Spalte erfolgte mit *Oder*. In der Stichwortsuche wurden Trunkierungen verwendet, um unterschiedliche Suffixe der Suchbegriffe zu erfassen. Um sowohl der britischen als auch der US-amerikanischen Schreibweise gerecht zu werden, wurden Wildcards eingesetzt. In der Thesaurussuche wurde keine Erweiterung (Explosion)

der Begriffe vorgenommen, sondern relevante Begriffe jeweils eigenständig aufgeführt.

Das Wortfeld *Krankheit* umfasste Formen der Suizidalität, Suizid und Selbstverletzung, einschließlich synonyme Begriffe. Das Wortfeld *Test* umfasste sowohl allgemeine Begrifflichkeiten für Tests, sowie ausgewählte kurze Screener für Suizidalität, die in einer Vorrecherche detektiert wurden (4, 63, 74, 87-89). Das Wortfeld *Diagnostische Genauigkeit* umfasste adäquate Maße, sowie den Überbegriff diagnostische Genauigkeit.

Tabelle 7: Stichwortsuche

Stichwortsuche		
<i>Krankheit</i>	<i>Test</i>	<i>Diagnostische Genauigkeit</i>
suicide*, self killing*, suicidal poisoning, self inflicted death, suicidal behavior*, suicid* risk*, suicid* ideation*, suicide attempt*, attempted suicide*, parasuicide*, self harm*	screen* trial*, screen*, questionnaire*, interview*, risk assessment*, risk analysis*, risk evaluation*, safety assessment*, self report *, diagnos*, diagnos* screen*, tool*, instrument*, suicidal ideation screen*, Paykel Suicide Item*, Suicide Behaviors Questionnaire, Suicide Ideation Scale, Suicidal Ideation Screening Questionnaire, Beck Depression Inventory, Hamilton Rating Scale for Depression, Beck Fast scan, Suicide risk screen, Columbia Suicide Severity Rating Scale, Suicide Behaviors Questionnaire AND Revised, Patient Safety Screener, Suizidalitätsitem aus dem Rasch-basierten Depressionsscreening, Self-Inflicted Injury Severity Form, SAD PERSONS scale, InterSePT Scale for Suicidal Thinking, Reasons for Living Inventory, Interpersonal Needs Questionnaire, Patient Health Questionnaire, P4, Symptom Driven Diagnostic System for Primary Care, Montgomery-Asberg Depression Rating Scale	valid*, reliab*, predictive value*, ROC curve, sensitiv*, specificity*

Tabelle 8: Thesaurussuche

Thesaurussuche		
<i>Krankheit</i>	<i>Test</i>	<i>Diagnostische Genauigkeit</i>
MEDLINE und Cochrane Library		
suicide/prevention and control, suicidal ideation, suicide attempted	risk assessment, diagnostic techniques and procedures, surveys and questionnaires, diagnostic self evaluation, self report, psychiatric status rating scales, interview as topic, interview, psychological, mass screening	sensitivity and specificity, ROC curve, reproducibility of results, diagnostic test, predictive value
EMBASE		
suicide/diagnosis prevention, suicidal behaviour, suicidal ideation, suicide attempt	questionnaire, interview, risk assessment, self report, diagnosis, psychological rating scale, psychologic assessment, psychological interview	diagnostic accuracy, diagnostic test, sensitivity and specificity, predictive value, diagnostic test accuracy study, receiver operating characteristic, validity, reliability
PsychINFO und PSYINDEX		
suicide, suicide prevention, suicidal ideation, attempted suicide, suicidology	questionnaires, interviews, risk assessment, screening tests, diagnosis, diagnostic interview schedules, psychodiagnostic interview, self-examination (medical), screening	test validity, test specificity, test sensitivity, test reliability

Die vollständige Suchstrategie befindet sich in Anhang G.

3.8 STUDIENSELEKTION

3.8.1 DEDUPLIKATION

Die Ergebnisse der elektronischen Literaturrecherche und der manuellen Recherche wurden nach EndNote X9 exportiert und dort gespeichert. Duplikate wurden entfernt. Dies erfolgte zunächst über eine automatisierte Funktion von Endnote, in der ein Vergleich der Felder *Jahr* und *Titel* stattfand. Anschließend wurde dies manuell überprüft.

3.8.2 TITEL- UND ABSTRACTSELEKTION

Für die Titel- und Abstractselektion wurde Rayyan, eine kostenfreie Web-Applikation des Qatar Computing Research Institutes für die Erstellung systematischer Übersichtsarbeiten, verwendet (90). Beide Rezensenten PD Dr. phil. Karoline Lukaschek (KL) und Milena Frank (MF) bewerteten den jeweiligen Artikel unabhängig voneinander anhand des Titels und Abstracts und kennzeichneten ihn als ein- oder ausgeschlossen. Nach Abschluss wurde die Verblindung der Rezensenten aufgehoben und sogenannte Konflikte (Unstimmigkeiten des Ein- oder Ausschlusses einer Studie) geklärt. Wenn keine Einigung der beiden Rezensenten stattfinden konnte, stand ein dritter Rezensent, Prof. Dr. med. Dipl.-Päd. Jochen Gensichen, MPH (JG) klärend zur Verfügung.

3.8.3 VOLLTEXTBESCHAFFUNG

In einem weiteren Schritt wurden die Volltexte der in der Titel- und Abstractselektion eingeschlossenen Artikel beschafft. Dies erfolgte über die Lizenzen der Universitätsbibliothek der Ludwig-Maximilians-Universität München, die wissenschaftlichen Suchmaschinen GoogleScholar, OAster und BASE, das Portal ProQuest, die kostenpflichtige Dokumentenlieferung von Subito und Anschreiben der Autoren auf der Plattform ResearchGate.

3.8.4 VOLLTEXTSELEKTION

Die Volltextselektion erfolgte anhand der festgelegten PICOS Kriterien. Ein Rezensent (MF) selektierte die Artikel anhand der Ein- und Ausschlusskriterien und vermerkte den Einschluss oder begründeten Ausschluss in einem vorbereiteten Microsoft Access Formular. Ein zweiter Rezensent (KL) überprüfte die Selektion. Unstimmigkeiten wurden im Konsensus gelöst. Für eine endgültige Entscheidung stand ein dritter Rezensent (JG) zur Verfügung.

3.9 DATENSAMMLUNG UND DATENEXTRAKTION

Relevante Daten wurden in einem Microsoft Access Formular erfasst. Ein Rezensent (MF) extrahierte die Daten und ein zweiter Rezensent (KL) überprüfte diese. Unstimmigkeiten wurden im Konsensus gelöst. Ein dritter Rezensent (JG) stand für eine Konfliktlösung zur Verfügung.

Die STARD Checkliste (91) diente als Vorlage zur Erstellung des Microsoft Access Formulars. Hierbei handelt es sich um eine Checkliste, die die wesentlichen Elemente diagnostischer Genauigkeitsstudien auflistet und Autoren diagnostischer Genauigkeitsstudien als Leitfaden dient (92).

Die STARD Checkliste befindet sich in Anhang G (91).

Wichtige Charakteristika stellten Studienpopulation, Setting, Ziel-Zustand, Indextest, Referenzstandard und Stichprobengröße dar (79).

3.10 STATISTISCHE METHODEN

Die Maße der diagnostischen Genauigkeit - Sensitivität, Spezifität, Negativer prädiktiver Vorhersagewert (NPV) und Positiver prädiktiver Vorhersagewert (PPV) - welche nicht im Originalartikel beschrieben wurden, wurden mittels ThresholdROC package version 2.7 in R version 3.6.0. berechnet.

Tabelle 9 gibt eine Darstellung der Vierfeldertafel eines diagnostischen Tests wieder.

Tabelle 9: Vierfeldertafel

		Referenzstandard	
		Positiv	Negativ
Indextest	Positiv	A Richtig positiv (RP)	B Falsch positiv (FP)
	Negativ	C Falsch negativ (FN)	D Richtig negativ (RN)

Hier gilt: Sensitivität = $\frac{A}{A+C}$; Spezifität = $\frac{D}{B+D}$; PPV = $\frac{A}{A+B}$; NPV = $\frac{D}{C+D}$

95% Konfidenzintervalle (KI) wurden, wenn diese nicht angegeben waren, anhand des Wilson-Intervalls berechnet.

Dr. rer. biol. hum., Dipl.-Biol. Kathrin Halfter (KH) prüfte die statistischen Berechnungen.

3.11 QUALITÄTBEWERTUNG

3.11.1 VERZERRUNGSRISSKO UND BEDENKEN BEZÜGLICH DER ANWENDBARKEIT

Das Verzerrungsrisiko und die Bedenken bezüglich der Anwendbarkeit wurden anhand des QUADAS-2 Tools bewertet. Hierbei handelt es sich um ein Tool zur Bewertung von Studien zur diagnostischen Genauigkeit, welches sich in vier Phasen gliedert. Nachdem in einem ersten Schritt die Forschungsfrage formuliert wird, wird im zweiten Schritt das Quadas-2 Tool für den spezifischen Review adaptiert. Anschließend werden Flussdiagramme für die jeweiligen Artikel erstellt. Final wird das Verzerrungsrisiko und die Bedenken bezüglich der Anwendbarkeit abgeschätzt. Die Einschätzung des Verzerrungsrisikos gliedert sich in vier Domänen. Es werden die Kategorien *Patientenselektion*, *Indextest*, *Referenzstandard* und *Flow und Timing* bewertet. Bezogen auf die Bedenken bezüglich der Anwendbarkeit werden ausschließlich die ersten drei Kategorien *Patientenselektion*, *Indextest* und *Referenzstandard* bewertet. Das Verzerrungsrisiko und Bedenken bezüglich der Anwendbarkeit werden als *hoch*, *gering* oder *unklar* eingestuft. (93)

Die QUADAS-2 Bewertung erfolgte anhand des von der University of Bristol zur Verfügung gestellten Quadas-2-Extraktionssheets (94) und wurde von den Rezensenten KL und MF unabhängig voneinander durchgeführt. Konflikte wurden im Konsensus gelöst. War eine Einigung nicht möglich, stand ein dritter Rezensent JG für eine Entscheidungsfindung zur Verfügung.

3.11.2 DIE GEWISSHEIT DER EVIDENZ (CERTAINTY OF EVIDENCE)

Die Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group stellt ein Bewertungssystem zur Verfügung, welches es ermöglicht, die Gewissheit der Evidenz für Ergebnisse systematischer Übersichtsarbeiten zu beurteilen (95). In dieser systematischen Übersichtsarbeit wurde GRADE für diagnostische Fragestellungen angewendet: *The GRADE approach for diagnostic tests and strategies* (96).

In der GRADEpro Software wurde in einem ersten Schritt jeweils eine diagnostische Fragestellung für den jeweiligen Indextest hinzugefügt. Dies erfolgte nach dem Schema: *Soll (Indextest) angewendet werden um (Zielproblem) bei (gesundheitliches Problem und/oder Population) zu diagnostizieren?* (97). Anschließend wurde eine Bewertung der unterschiedlichen Kategorien *Studiendesign, Risiko für Bias, Indirektheit, Inkonsistenz, Fehlende Genauigkeit* und *Andere Faktoren* für die eingeschlossenen Indextests durchgeführt. Anhand dieser Einschätzung erstellte die Software GRADEpro eine Bewertung der Gewissheit der Evidenz als *sehr gering, gering, moderat* oder *hoch*; siehe *Tabelle 10*.

Auf Empfehlung eines Online-Trainings der Cochrane Collaboration und der McMaster University of Health Sciences zur Kategorie *Inkonsistenz*, wurden Screener, die in nur einem Artikel untersucht wurden, in der Bewertung der *Inkonsistenz* nicht herabgestuft (98).

Tabelle 10: Klassifizierung der Gewissheit der Evidenz. Übersetzt nach Tabelle „Information Certainty of evidence“ (97)

Klassifizierung	Definition
Hoch	Es ist sehr unwahrscheinlich, dass weitere Untersuchungen unser Vertrauen in die Einschätzung des Effektes verändern.
Moderat	Weitere Untersuchungen dürften einen wichtigen Einfluss auf unser Vertrauen in die Einschätzung des Effekts haben und können die Einschätzung ändern.
Gering	Weitere Untersuchungen werden sehr wahrscheinlich einen wichtigen Einfluss auf unser Vertrauen in die Einschätzung des Effekts haben und die Einschätzung wahrscheinlich ändern.
Sehr gering	Die Einschätzung des Effekts ist sehr unsicher.

3.12 DATENSYNTHESE

Aufgrund der großen Heterogenität der Screening Instrumente wurde eine qualitative Synthese der Daten einer quantitativen Synthese vorgezogen (99). Sensitivität und Spezifität sowie deren 95% Konfidenzintervalle wurden in einem Forest Plot dargestellt.

4 ERGEBNISSE

4.1 STUDIENSELEKTION

Die Literaturrecherche ergab eine Gesamtstudienzahl von $N = 12\,460$ Treffern. Diese entstammten zum einen der elektronischen Datenbanksuche ($n = 12\,403$), zum anderen der manuellen Recherche ($n = 57$). Nach der Überprüfung auf Duplikate wurden Titel und Abstract der verbliebenen $6\,870$ Artikel bewertet. Hierbei wurden $6\,799$ Treffer ausgeschlossen und die restlichen 71 Artikel anhand ihrer Volltexte beurteilt. Letztendlich wurden 64 Artikel begründet ausgeschlossen. Hauptursache für die Ausschlüsse stellten die Kategorie *Population* dar ($n = 34$), gefolgt von der Kategorie *Diagnostische Genauigkeit* ($n = 8$). In *Abbildung 2* ist eine Übersicht der Studienselektion ersichtlich. Somit wurden sieben Studien identifiziert, die in die systematische Übersichtsarbeit eingeflossen sind. (83)

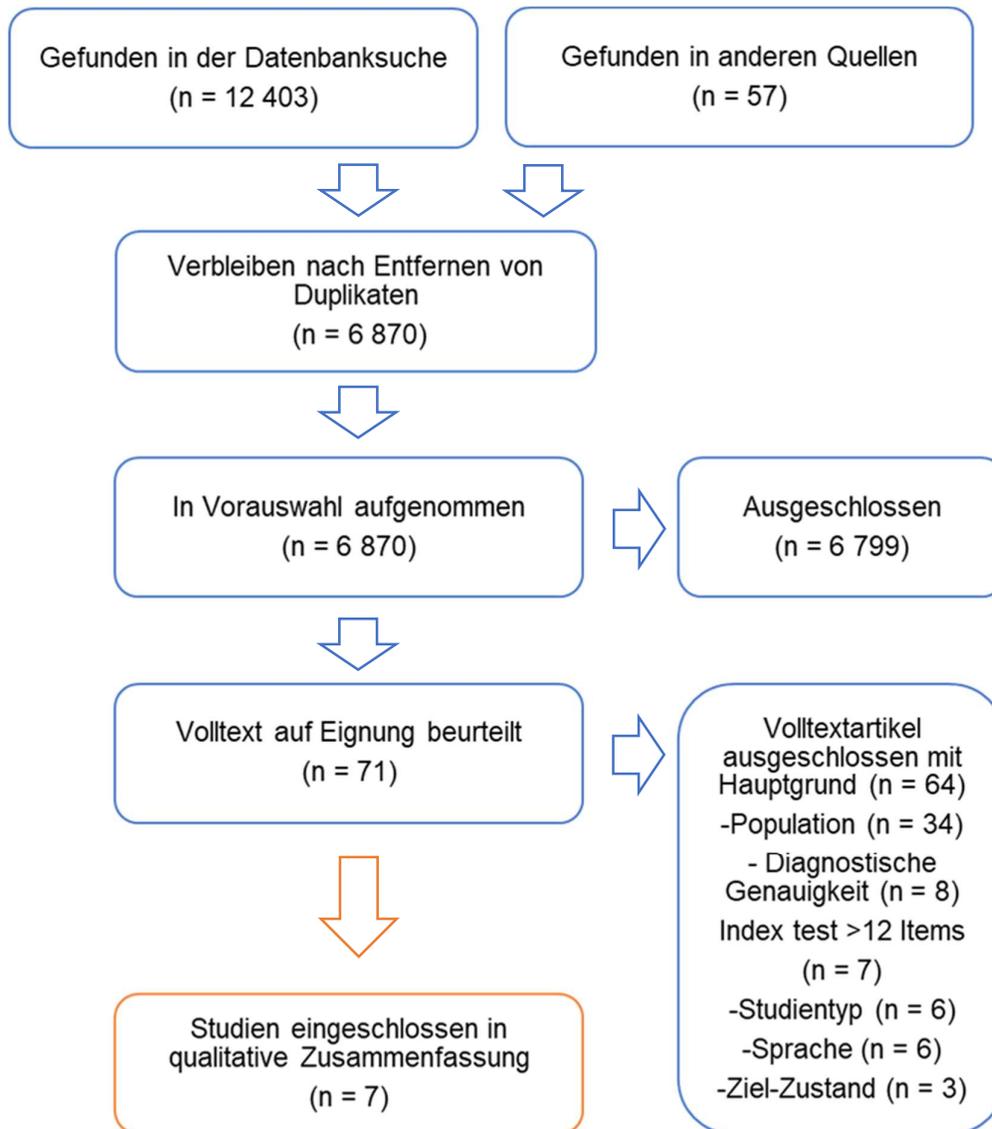


Abbildung 2: PRISMA Flussdiagramm nach Vorlage von Moher et al. (2009) (100)

4.2 CHARAKTERISTIKA DER EINGESCHLOSSENEN STUDIEN

4.2.1 ALLGEMEINES

Folgende sieben Studien wurden in den systematischen Review eingeschlossen:

- *Identifying Suicidal Ideation in General Medical Patients*, Cooper-Patrick et al., 1994, JAMA (101)

- *Suicidal Ideation in Primary Care*, Olfson et al., 1996, Journal of general internal medicine (102)
- *Patient Health Questionnaire Depression Scale as a Suicide Screening Instrument in Depressed Primary Care Patients: A Cross-Sectional Study*, Uebelacker et al., 2011, The primary care companion for CNS disorders (103)
- *Suicidal Ideation - The Clinical Utility of the K10*, Chamberlain et al., 2009, Crisis (104)
- *Single-Item Measurement of Suicidal Behaviors: Validity and Consequences of Misclassification*, Millner et al., 2015, PloS one (105)
- *The Suicidal Ideation Attributes Scale (SIDAS): Community-Based Validation Study of a New Scale for the Measurement of Suicidal Ideation*, Van Spijker et al., 2014, Suicide & life-threatening behavior (106)
- *Development and Validation of Static and Adaptive Screeners to Assess Suicidal Thoughts and Behavior*, Calear et al. 2020 Suicide & life-threatening behavior (107)

Die Studien untersuchten jeweils unterschiedliche Indextests und wurden jeweils gegen einen unterschiedlichen Referenzstandard evaluiert. Drei der Studien untersuchten Suizidalität in der Allgemeinmedizin (101-103) und vier der Studien in der Allgemeinbevölkerung (104-107).

Die Studien wurden entweder in Australien (104, 106, 107) oder den Vereinigten Staaten von Amerika (101-103, 105) durchgeführt. Der Studienzeitraum erstreckte sich über mehrere Jahrzehnte, wobei die älteste Studie im Jahr 1994 (101) und die aktuellste im Jahr 2020 publiziert wurde (107). Bei allen eingeschlossenen Studien handelte es sich um Querschnittsstudien. Untersuchte Ziel-Zustände waren Suizidgedanken (vier Mal) (101, 102, 104, 105), Suizidversuch (drei Mal) (105-107), Suizidplan (zwei Mal) (102, 106) und Suizidalität im Allgemeinen (einmal) (103). Die Studien unterschieden sich hinsichtlich der Stichprobengröße. Die Teilnehmeranzahl der kleinsten Studie betrug $n = 166$ (103), die Teilnehmeranzahl der größten Studie betrug $n = 12\,884$ (104). Die Studie von *Uebelacker et al. (2011)* untersuchte als einzige Studie Risikopatienten (Patienten mit Depression) (103). (83)

Tabelle 11 gibt einen Überblick über die Charakteristika der eingeschlossenen Studien
(Die Tabelle wurde von MF erstellt und ist Teil der Publikation (83)).

Tabelle 11: Charakteristika der eingeschlossenen Studien

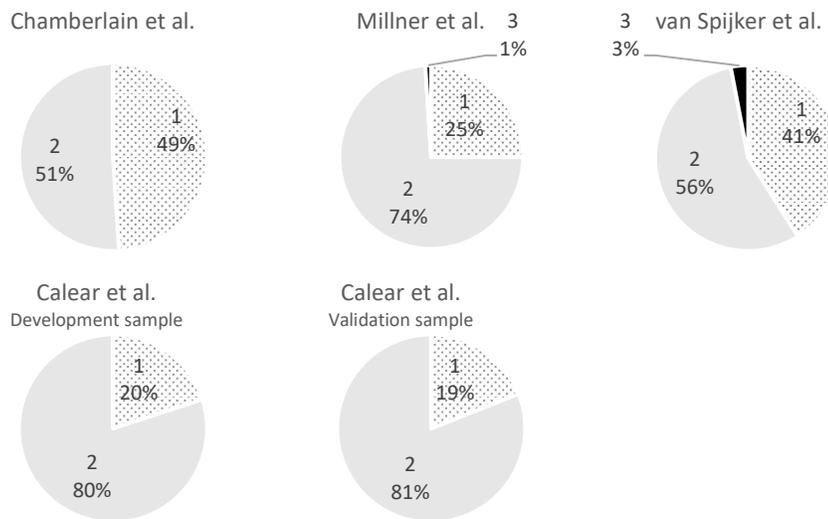
Autor, Jahr, Land	Population, Alter	Setting	Ziel-Zustand	Indextest, Anzahl der Items, Selbstbeurteilung/Interview	Referenzstandard	Stichprobengröße
Calear et al., 2020, Australien	Allgemeinbevölkerung, Teilnehmer >18 Jahre	Onlineumfrage	Suizidversuch	1. Rapid Measurement Toolkit for Suicidal Behaviors Questionnaire- Revised, Fünf Items, Selbstbeurteilung 2. Suicidal Behaviors Questionnaire- Revised, Vier Items, Selbstbeurteilung	Suicidality Item Bank	1. Development sample: n = 3 175 2. Validation sample: n = 3 577
Chamberlain et al., 2009, Australien	Allgemeinbevölkerung, Teilnehmer > 16 Jahre	Telefoninterview	Suizidgedanken	Kessler Psychological Distress Scale, Zehn Items, Interview	Subskala des General Health Questionnaire 28	n = 12 884
Millner et al., 2015, USA	Allgemeinbevölkerung, Teilnehmer ≥ 18 Jahre	Onlineumfrage	Suizidgedanken, Suizidversuch	1. Gate question suicide attempt, Ein Item, Selbstbeurteilung 2. Gate question suicide ideation, Ein Item, Selbstbeurteilung	1. Kodierung als Suizidversuch 2. Follow-up questions suicide ideation	n = 1 231 Ziel-Zustand Suizidversuch n = 1 570 Ziel-Zustand Suizidgedanken
Van Spijker et al., 2014, Australien	Allgemeinbevölkerung, Teilnehmer ≥ 18 Jahre	Onlineumfrage	Suizidplan, Suizidversuch	Suicidal Ideation Attributes Scale, Fünf Items, Selbstbeurteilung	Subskala des Columbia-Suicide Severity Rating Scale	n = 1 352
Cooper-Patrick et al., 1994, USA	Allgemeinmedizin, Teilnehmer ≥ 18 Jahre	Haushaltsumfrage	Suizidgedanken	Suicidal Ideation Screening Questionnaire, Vier Items, Interview	Subskala des Diagnostic Interview Schedule	n = 6 041
Olfson et al., 1996, USA	Allgemeinmedizin, Teilnehmer ≥ 18 Jahre	Allgemeinmedizinische Sprechstunde	Suizidgedanken, Suizidplan	1. Feeling suicidal, Ein Item, Selbstbeurteilung 2. Wishing you were dead, Ein Item, Selbstbeurteilung 3. Thoughts of death, Ein Item, Selbstbeurteilung 4. Feeling suicidal, Ein Item, Selbstbeurteilung	1-3. Frage eines strukturierten Interviews zu Suizidplan 4. Strukturiertes Interview mit Fragen zu aktueller und vorangehender Suizidalität	n = 1 001
Uebelacker et al., 2011, USA	Allgemeinmedizin, Teilnehmer ≥ 18 Jahre	Allgemeinmedizinische Sprechstunde	Suizidalität	Patient Health Questionnaire 9 - Item, Ein Item, Selbstbeurteilung	Subskala des Mood module of the Structured Clinical Interview for DSM-IV Axis I Disorders	n = 166

4.2.2 TEILNEHMER DER STUDIEN

4.2.2.1 GESCHLECHTERVERTEILUNG

Die Geschlechterverteilung der jeweiligen Studien wird in den folgenden *Abbildungen 3 und 4* veranschaulicht. Mit Ausnahme der Studie von *Chamberlain et al. (2009)* (104) war der Anteil an Frauen im Gegensatz zu dem Anteil an Männern deutlich höher.

Chamberlain et al. (2009) (104) verwies in seinem Artikel auf die Studie von *Taylor et al. (2007)* (108). Hier wurde jedoch nur ein Teil der Studienpopulation beschrieben (n = 5 272). Die Verteilung der Geschlechter von *Taylor et al. (2007)* (108) wurde stellvertretend für *Chamberlain et al. (2009)* (104) im folgenden Diagramm dargestellt. *Millner et al. (2015)* (105) und *Olfson et al. (1996)* (102) stellten die Geschlechterverteilung für die gesamte Studienpopulation dar. Diese wurde auch im folgenden Diagramm dargestellt. *Calear et al. (2020)* (107) gab die Charakteristika für das Development sample und das Validation sample getrennt an.



1=Männer, 2=Frauen 3=Andere/Unbekannt

Abbildung 3: Geschlechterverteilung in den Studien der Allgemeinbevölkerung

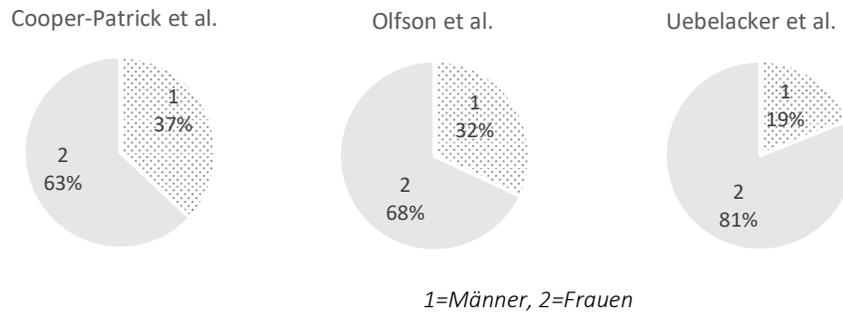


Abbildung 4: Geschlechterverteilung in den Studien der Allgemeinmedizin

4.2.2.2 EINGESCHLOSSENE ALTERSSTUFEN

Die Altersverteilung wird in folgender *Abbildung 5* visualisiert. Die Studien schlossen jeweils eine große Altersspanne der Studienteilnehmer ein.

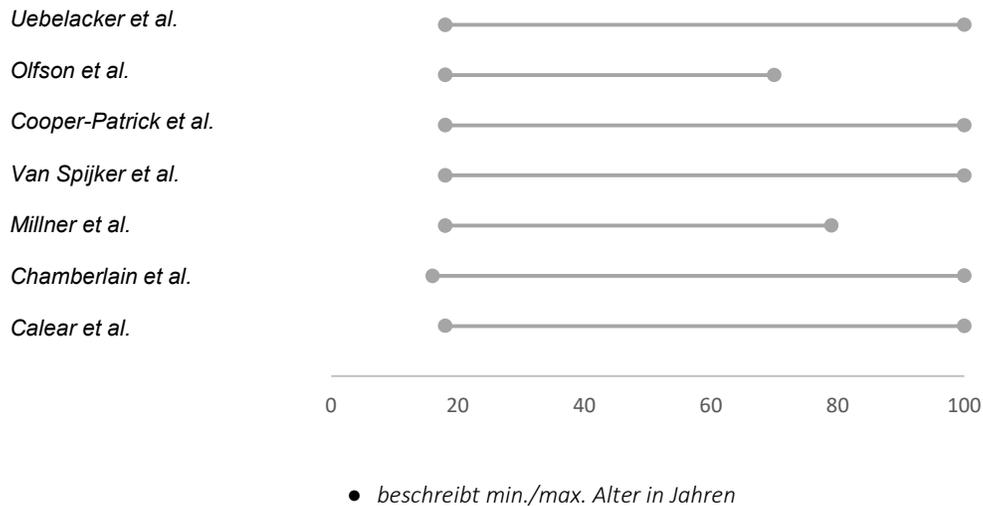


Abbildung 5: Alter der Studienteilnehmer

4.2.3 INDEXTEST UND REFERENZSTANDARD

4.2.3.1 INDEXTEST

Alle eingeschlossenen Studien untersuchten jeweils einen anderen Indextest. Einige Autoren analysierten mehrere Indextests in einem Artikel (102, 105, 107) oder prüften

denselben Indextest gegen unterschiedliche Referenzstandards (102, 106). Unterschiedliche Cut-Off-Werte wurden von *van Spijker et al. (2014)* (106) beschrieben, sowie für *Chamberlain et al. (2009)* (104) berechnet.

Calear et al. (2020) (107) entwickelte den *Rapid Measurement Toolkit for Suicidal-Static Screener* (RMTS-S) und evaluierte diesen in der Allgemeinbevölkerung bezüglich dem Ziel-Zustand Suizidversuch. Der Fragebogen enthält fünf Items, die mit einer fünfstufigen Likert-Skala bewertet werden. Zudem wurde der in bereits vielen Studien verwendete *Suicidal Behaviors Questionnaire – Revised* (SBQ-R) eingesetzt. Dabei handelt es sich um einen Selbstbeurteilungsbogen mit vier Items. Anschließend wurden die Ergebnisse beider Indextests bezüglich ihrer diagnostischen Genauigkeit verglichen.

Chamberlain et al. (2009) (104) untersuchte die Fähigkeit des *Kessler Psychological Distress Scale* (K10) Suizidgedanken in der Allgemeinbevölkerung zu erkennen. Der K10 wurde von *Kessler und Mroczek (1994)* entwickelt (109) zitiert nach (104). Es handelt sich um einen Fragebogen mit zehn Items, der ursprünglich für die Erkennung von Symptomen der Angst und Depression in der Allgemeinbevölkerung entwickelt wurde. Die Antworten konnten mit einer fünfstufigen Likert-Skala angegeben werden. Als Cut-Off wurden die Werte 22 (hoch) und 30 (sehr hoch) gewählt. Der K10 konfrontiert die Studienteilnehmer nicht direkt mit Suizidalität.

Millner et al. (2015) (105) untersuchte zwei Indextests mit nur jeweils einem Item. Um Suizidgedanken in der Allgemeinbevölkerung zu detektieren, stellte der Autor die Frage, ob der Studienteilnehmer jemals ernsthafte Suizidgedanken hatte. Um Suizidversuche zu detektieren, stellte der Autor eine direkte Frage zu vorangegangenen Suizidversuchen. Die sogenannte *Gate question suicide attempt* und *Gate question suicide ideation* konnte jeweils mit *Ja* oder *Nein* beantwortet werden.

Van Spijker et al. (2014) (106) entwickelte den *Suicidal Ideation Attributes Scale* (SIDAS) für die Detektion eines Suizidplans oder Suizidversuchs. Es handelt sich um einen Fragebogen mit fünf Items, der die Häufigkeit, Kontrollierbarkeit, Belastung und Beeinträchtigung des Alltags von Suizidgedanken sowie die Gefahr eines Suizidversuchs abfragt. Die unterschiedlichen Items konnten jeweils auf einer zehn-Punkte-

Skala beantwortet werden. Im Artikel beschreibt der Autor einen Cut-Off-Wert von eins, sowie 21 für beide Ziel-Zustände.

Cooper-Patrick et al. (1994) (101) entwickelte den *Suicidal Ideation Screening Questionnaire (SIS-Q)* um Suizidgedanken in der Allgemeinmedizin zu erkennen. Es handelt sich um einen Fragebogen mit vier Items, der vier unterschiedliche Symptome abfragt: Schlafstörung, Stimmungsstörung, Hoffnungslosigkeit und Schuldgefühle/Wertlosigkeit. Auch in diesem Fragebogen wurden die Probanden nicht direkt mit Suizidalität konfrontiert. Antwortmöglichkeiten beinhalteten *Ja* oder *Nein*. Der Autor wählte einen Cut-Off-Wert von eins.

Olfson et al. (1996) (102) prüfte die Fähigkeit der Indextests mit jeweils nur einem Item - *Feeling suicidal*, *Wishing you were dead* und *Thoughts of death* - Patienten mit einem Suizidplan in der Allgemeinmedizin herauszufiltern. Den Indextest *Feeling suicidal* verwendete der Autor zudem, um Patienten mit Suizidgedanken zu identifizieren. Antwortmöglichkeiten beinhalteten *Ja* oder *Nein*.

Uebelacker et al. (2011) (103) prüfte ob *Item 9* des *Patient Health Questionnaire - 9 (PHQ-9)* Suizidalität in der Allgemeinmedizin detektieren könnte. Der *PHQ-9* (110) wurde bereits für das Screening von Depression vielfach im klinischen Alltag eingesetzt. *Item 9* setzt sich mit der Frage nach Suizidgedanken und selbstverletzendem Verhalten auseinander. Als Antwortmöglichkeiten standen *Not at all*, *Several days*, *More than half the days* und *Nearly every day* zur Verfügung. Ab dem Cut-Off-Wert *Several days* wurde Suizidalität angenommen.

4.2.3.2 REFERENZSTANDARD

Die Autoren der eingeschlossenen Studien verwendeten jeweils unterschiedliche Referenzstandards. *Olfson et al. (1996)* (102) und *van Spijker et al. (2014)* (106) untersuchten jeweils einen Indextest im Vergleich zu zwei unterschiedlichen Ziel-Zuständen eines Referenzstandards.

Calear et al. (2020) (107) entwickelte in einem aufwändigen mehrstufigen Prozess die *Suicidality Item Bank*. Die Items wurden aus einer systematischen Literaturrecherche für bereits bestehende Fragebögen zu Suizidalität extrahiert. Nach anschließender Standardisierung hinsichtlich Fragenstruktur und darauffolgender Begutachtung durch Forscher, Kliniker und Betroffene wurden letztlich 23 Items ausgewählt. Die Items beinhalten Fragen bezüglich Suizidgedanken, -plänen und -versuchen. Diese konnten auf einer fünfstufigen Likert-Skala beantwortet werden: *Never, Rarely, Sometimes, Often und Always*.

Chamberlain et al. (2009) (104) verwendete eine Subskala des *General Health Questionnaire - 28 (GHQ-28)* (111) als Referenzstandard um Suizidgedanken zu erfassen (D4, D5, D6 und D7). Die Subskala wurde bereits von *Watson et al. (2001)* (112) beschrieben und beinhaltet vier unterschiedliche Aspekte von Suizidgedanken (108). Binäre Antwortmöglichkeiten wurden angeboten. So erhielten die Studienteilnehmer null bis vier Punkte. Ab einem Cut-Off-Wert von einem Punkt wurden Suizidgedanken angenommen.

Millner et al. (2015) (105) validierte die *Gate question suicide attempt* mit einer aufwändigen Kodierung, welche im Artikel genau beschrieben wurde. Für die *Gate question suicide ideation* verwendete *Millner et al. (2015)* (105) die *Follow-up questions suicide ideation* als Referenzstandard. Diese Fragen steigern sich von generellen Gedanken über den Tod über passive Suizidgedanken, hin zu aktiven Suizidgedanken. Als Cut-Off-Wert wurde nach Rücksprache mit dem Letztautor Prof. Mathew Nock, Harvard University, aktive Suizidgedanken gewählt. Mögliche Antwortoptionen waren *Ja, Nein* oder *Keine Antwort*.

Van Spijker et al. (2014) (106) verwendete eine Teilskala des *Columbia - Suicide Severity Rating Scale (C-SSRS)* (113) als Referenzstandard. Diese beinhalteten Fragen zu Suizidplan und Suizidversuch innerhalb des letzten Jahres. Der Autor validierte den Indextest *SIDAS* zum einen gegen Fragen zu Suizidplan und zum anderen gegen eine Frage zu Suizidversuch. Antwortmöglichkeiten stellten fünf- bis sechs-Punkte-Skalen dar. Der verwendete Cut-Off-Wert wurde nicht berichtet.

Cooper-Patrick et al. (1994) (101) verwendete vier Fragen des *Diagnostic Interview Schedule (DIS)*, ein strukturiertes Interview, als Referenzstandard (114). Zu Beginn wurde nach Gedanken an Tod generell gefragt, im Verlauf nach konkreten Suizidgedanken und Suizidversuchen (101). Der Cut-Off-Wert wurde durch ein bejahen konkreter Suizidgedanken im letzten Jahr definiert.

Olfson et al. (1996) (102) verwendete als Referenzstandard ein *strukturiertes Interview* (nicht genauer beschrieben) zu vorangegangener und aktueller Suizidalität. Dieses teilte der Autor in der Auswertung zum einen in die Frage nach einem Suizidplan und zum anderen in die Frage nach Suizidgedanken auf. Ausschließlich der Indextest *Feeling suicidal* wurde bezüglich des Ziel-Zustandes Suizidgedanken ausgewertet.

Uebelacker et al. (2011) (103) verwendete eine Subskala des *Mood Module of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCD-I mood module)* (115). Die Subskala beurteilt die Aspekte *Gedanken an den Tod, Suizidgedanken, Suizidplan* und *Suizidversuch*. Als Cut-Off-Wert wurden *wiederkehrende Gedanken an den Tod* eingesetzt.

4.3 QUALITÄTBEWERTUNG

Einen Überblick über die Ergebnisse der Bewertung des Verzerrungsrisikos und Bedenken bezüglich der Anwendbarkeit geben *Tabelle 12* und *Abbildung 6* (*Tabelle 12* wurde von MF erstellt und ist Teil der Publikation (83)). Die Qualitätsbewertung des *RMTS-S Development samples* (107) wird aus Gründen der Übersichtlichkeit nicht dargestellt.

Tabelle 12: QUADAS-2 Tool nach Vorlage der tabellarischen Darstellung der University of Bristol (116)

Studie	Verzerrungsrisiko				Bedenken bezüglich der Anwendbarkeit		
	Patienten-selektion	IT	RS	Flow und Timing	Patienten-selektion	IT	RS
Calear et al.	Hoch	Gering	Gering	Gering	Gering	Gering	Gering
Chamberlain et al.	Gering	Gering	Unklar	Gering	Gering	Gering	Gering
Millner et al.	Unklar	Gering	Gering	Unklar	Unklar	Gering	Gering
Van Spijker et al.	Hoch	Gering	Unklar	Unklar	Unklar	Gering	Gering
Cooper-Patrick et al.	Unklar	Unklar	Gering	Gering	Gering	Gering	Gering
Olfson et al.	Gering	Gering	Unklar	Gering	Gering	Gering	Gering
Uebelacker et al.	Hoch	Gering	Unklar	Hoch	Gering	Gering	Gering

IT = Indextest; RS = Referenzstandard

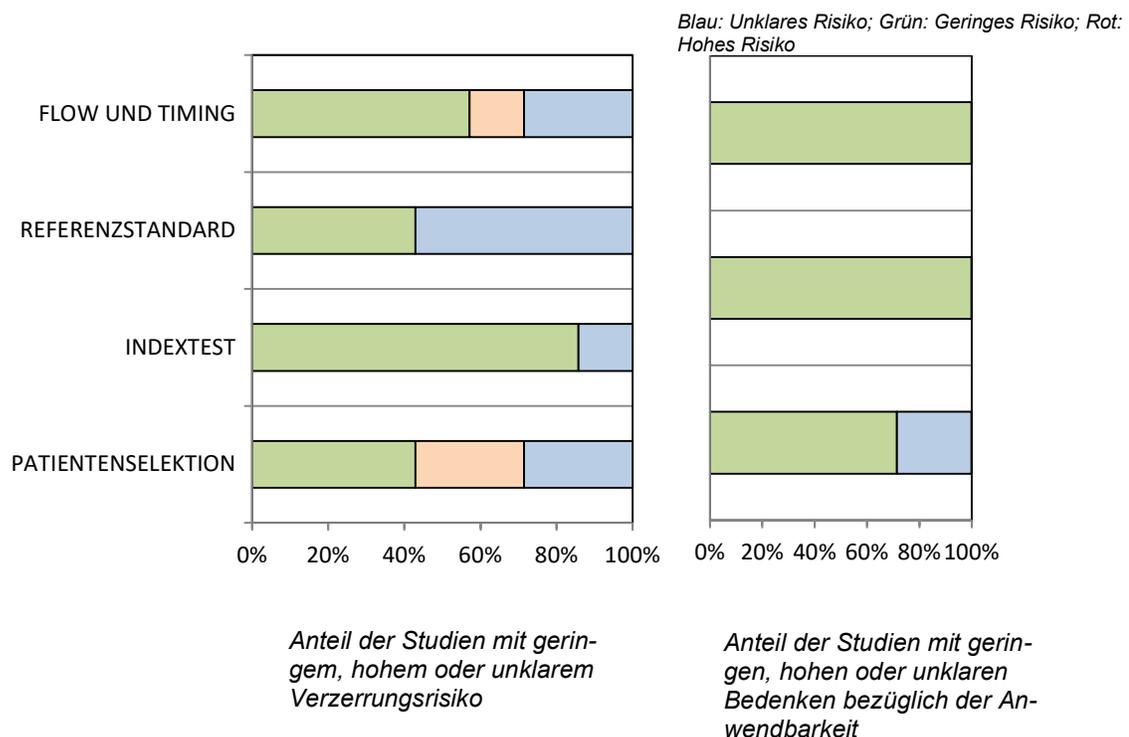


Abbildung 6: QUADAS-2 Domäne nach Vorlage der graphischen Darstellung der University of Bristol (117)

4.3.1 VERZERRUNGSRISIKO

In der Kategorie Patientenselektion wies das Verzerrungsrisiko in 29% der Fälle ein *geringes*, in 43% ein *hohes* und in 29% ein *unklares* Risiko auf. Das Risiko bei *Uebelacker et al. (2011)* (103), *van Spijker et al. (2014)* (106) und *Calear et al. (2020)* (107) wurde als *hoch* eingeschätzt, da die Rekrutierung der Teilnehmer einer willkürlichen Stichprobe entsprach. Unklarheit bestand bezüglich der Bewertung der Kategorie *Patientenselektion* bei *Millner et al. (2015)* (105). In dieser Studie wurden die Teilnehmer mittels Anzeigenwebsite rekrutiert, und zwei unterschiedliche Titel gewählt, um sowohl suizidale als auch nicht-suizidale Patienten anzusprechen. Auch bei *Cooper-Patrick et al. (1994)* (101) wurde die Patientenselektion als *unklar* eingestuft.

Die Kategorie Indextest wies in 86% der Fälle ein *geringes* und in 14% ein *unklares* Verzerrungsrisiko auf. Das Verzerrungsrisiko der Kategorie Indextest von *Cooper-Patrick et al. (1994)* (101) wurde als *unklar* bewertet, da es sowohl Unklarheiten bezüglich eines vorab festgelegten Schwellenwertes als auch der Interpretation des Indextests ohne die Ergebnisse des Referenzstandards gab.

Das Verzerrungsrisiko des Referenzstandards wurde in 43% als *gering*, sowie in 57% als *unklar* eingestuft. Unklar war, ob der Referenzstandard ohne Kenntnis der Ergebnisse des Indextests ausgewertet wurde (Verblindung).

In der Kategorie Flow und Timing wurden 57% der Studien mit einem *geringen* Verzerrungsrisiko beurteilt, 29% als *unklar* und 14% als *hoch*. Das Verzerrungsrisiko in der Kategorie Flow und Timing wurde in der Studie von *Übelacker et al. (2011)* (103) als *hoch* eingestuft, da hier eine mittlere Zeitspanne von 9 Tagen zwischen der Erhebung des Indextests und des Referenzstandards lag. Unklarheiten bestanden bezüglich des Zeitintervalls zwischen den beiden Tests bei *Millner et al. (2015)* (105) und *van Spijker et al. (2014)* (106). Das Zeitintervall wurde in diesen Studien nicht berichtet.

4.3.2 BEDENKEN BEZÜGLICH DER ANWENDBARKEIT

Bedenken bezüglich der Anwendbarkeit gab es in der Kategorie Patientenselektion. Hier fiel zweimal die Bewertung mit *unklar* aus. Zum einen wurde dieses Urteil in der Studie von *van Spijker et al. (2014)* (106) gefällt, der die Teilnehmer der Studie in einem sozialen Netzwerk rekrutierte. Zum anderen wurde dieses Urteil in der Studie von *Millner et al. (2015)* (105) gefällt, hier wurden die Teilnehmer über eine Anzeigenwebseite rekrutiert. Die restlichen Kategorien (Indextest und Referenzstandard) wiesen *geringe* Bedenken bezüglich der Anwendbarkeit auf.

4.3.3 GRAD DER ÜBEREINSTIMMUNG

Die Bewertung des Verzerrungsrisikos und Bedenken bezüglich der Anwendbarkeit der Rezensenten KL und MF stimmte in 84% der Fälle überein. Eine Einigung der restlichen 16% fand im Diskurs statt.

4.4 ERGEBNISSE DER EINZELNEN STUDIEN

Einen Überblick über die Ergebnisse der einzelnen Studien gibt *Tabelle 13 und Tabelle 14* (Die Tabellen wurde von MF erstellt und sind Teil der Publikation (83)).

In den Studien der Allgemeinbevölkerung betrug die Prävalenz von Suizidalität 4% (107) bis 75% (105). In der Studienpopulation Allgemeinmedizin ohne Vortest betrug die Prävalenz 1%-3% (101, 102). In der Studie mit einem Vortest für Depression betrug die Prävalenz von Suizidalität 35% (103).

Die Sensitivität der betrachteten Indextests variierte zwischen 26% und 100%, die Spezifität zwischen 64% und 99%. Auch der PPV wies eine große Bandbreite auf. Diese reichte von 6% (102) bis 91% (105). Der NPV wies Werte von 83% (103) bis 100% auf.

Die Gewissheit der Evidenz wurde in zwei Studien *hoch* (102, 104), vier Studien als *moderat* (101, 103, 105, 107) und einer Studie als *gering* (106) eingestuft. (83)

Tabelle 13: Die Ergebnisse der einzelnen Studien in der Allgemeinbevölkerung

Autor	Indextest	Referenzstandard	n	Prävalenz	RP n =	FP n =	RN n =	FN n =	Sensitivität (95% KI)	Spezifität (95% KI)	PPV (95% KI)	NPV (95% KI)	Gewissheit der Evidenz
Calear et al.	RMTS-S Development sample	Suicidality Item Bank	3 175	0,04	120	516	2 519	20	0,86 (0,79-0,91)	0,83 (0,82-0,84)	0,19 (0,16-0,22)	0,99 (0,99-1,00)	moderat
Calear et al.	RMTS-S Validation sample	Suicidality Item Bank	3 577	0,04	124	1 030	2 404	19	0,87 (0,80-0,92)	0,70 (0,68-0,72)	0,11 (0,09-0,13)	0,99 (0,99-1,00)	moderat
Calear et al.	SBQ-R	Suicidality Item Bank	3 175	0,04	115	516	2 519	25	0,82 (0,75-0,88)	0,83 (0,83-0,83)	0,18 (0,17-0,19)	0,99 (0,99-0,99)	moderat
Chamberlain et al.	K10 Cut-Off 30	GHQ-28 Subskala	12 069	0,05	159	146	11 311	453	0,26 (0,23-0,30)	0,99 (0,98-0,99)	0,52 (0,46-0,58)	0,96 (0,96-0,96)	hoch
Chamberlain et al.	K10 Cut-Off 22	GHQ-28 Subskala	12 069	0,05	356	821	10 636	256	0,58 (0,54-0,62)	0,93 (0,92-0,93)	0,30 (0,28-0,33)	0,98 (0,97-0,98)	hoch
Millner et al.	Gate question suicide attempt	Kodierung als Suizidversuch	1 231	0,37	401	48	732	50	0,89 (0,86-0,92)	0,94 (0,92-0,95)	0,89 (0,86-0,92)	0,94 (0,92-0,95)	moderat
Millner et al.	Gate question suicide ideation	Follow up questions for suicide ideation	1 570	0,75	1 129	109	286	46	0,96 (0,95-0,97)	0,72 (0,68-0,77)	0,91 (0,89-0,93)	0,86 (0,82-0,90)	moderat
Van Spijker et al.	SIDAS Cut-Off 1	C-SSRS-Suizidplan	1 352	0,16	188	372	760	32	0,86 (0,80-0,90)	0,67 (0,64-0,70)	0,34 (0,30-0,38)	0,96 (0,94-0,97)	gering
Van Spijker et al.	SIDAS Cut-Off 1	C-SSRS-Suizidversuch	1 352	0,11	121	439	769	23	0,84 (0,77-0,89)	0,64 (0,61-0,66)	0,22 (0,18-0,25)	0,97 (0,96-0,98)	gering
Van Spijker et al.	SIDAS Cut-Off 21	C-SSRS-Suizidplan	1 352	0,16	87	47	1 058	133	0,40 (0,33-0,46)	0,96 (0,94-0,97)	0,65 (0,56-0,72)	0,89 (0,87-0,91)	gering
Van Spijker et al.	SIDAS Cut-Off 21	C-SSRS-Suizidversuch	1 352	0,11	72	62	1 146	72	0,50 (0,42-0,58)	0,95 (0,93-0,96)	0,54 (0,45-0,62)	0,94 (0,93-0,95)	gering

Tabelle 14: Die Ergebnisse der einzelnen Studien in der Allgemeinmedizin

Autor	Indextest	Referenzstandard	n	Prävalenz	RP n =	FP n =	RN n =	FN n =	Sensitivität (95% KI)	Spezifität (95% KI)	PPV (95% KI)	NPV (95% KI)	Gewissheit der Evidenz
Cooper-Patrick et al.	SIS-Q Cut-Off 1	Subskala DIS	6 041	0,03	130	1 470	4 417	24	0,84 (0,77-0,90)	0,75 (0,74-0,76)	0,08 (0,07-0,10)	0,99 (0,99-1,00)	moderat
Olfson et al.	Feeling suicidal	Strukturiertes Interview- Suizidge-danken	1 001	0,03	25	8	964	4	0,86 (0,67-0,95)	0,99 (0,98-0,99)	0,76 (0,57-0,88)	1,00 (0,99-1,00)	hoch
Olfson et al.	Feeling suicidal	Strukturiertes Interview - Suizidplan	1 001	0,01	10	23	966	2	0,83 (0,62-1)	0,98 (0,97-0,99)	0,30 (0,15-0,46)	1,00 (-)	hoch
Olfson et al.	Thoughts of death	Strukturiertes Interview	1 001	0,01	12	190	799	0	1,00 (-)	0,81 (0,79-0,84)	0,06 (0,03-0,09)	1,00 (-)	hoch
Olfson et al.	Wishing you were dead	Strukturiertes Interview	1 001	0,01	11	68	921	1	0,92 (0,76-1)	0,93 (0,92-0,95)	0,14 (0,06-0,22)	1,00 (-)	hoch
Uebelacker et al.	PHQ-9, Item 9	Subskala SCID-I	166	0,35	40	17	91	18	0,69 (0,55-0,80)	0,84 (0,76-0,90)	0,70 (0,56-0,81)	0,83 (0,75-0,90)	moderat

4.5 GRUPPIERUNG DER ERGEBNISSE NACH DEM ZIEL-ZUSTAND

Im Folgenden wird eine Gruppierung nach dem Ziel-Zustand sowie ein Überblick über die jeweilige Sensitivität und Spezifität und das GRADE Rating gegeben. Zudem wird angegeben, welche Werte der diagnostischen Genauigkeit selbst berechnet und welche den Originalartikeln entnommen wurden. (Die folgenden Abschnitte 4.5.1 - 4.5.4 wurden von MF erstellt und sind Teil der Publikation (83)).

4.5.1 ZIEL-ZUSTAND SUIZIDGEDANKEN

Chamberlain et al. (2009) (104) evaluierte den *K10* hinsichtlich dem Ziel-Zustand Suizidgedanken. Für den Cut-Off-Wert *sehr hoch (30 Punkte)* wurde eine Sensitivität von 26% und Spezifität von 99% berechnet. Für den Cut-Off-Wert *hoch (22 Punkte)* wurde eine Sensitivität von 58% und eine Spezifität von 93% berechnet. Die Gewissheit der Evidenz wurde mit *hoch* eingestuft.

Millner et al. (2015) (105) untersuchte die *Gate question suicide ideation*. Nach Rücksprache mit dem Letztautor Prof. Mathew Nock, Harvard University, wurde *Maybe I should kill myself* als Cut-Off-Wert des Referenzstandards gewählt und eine Sensitivität von 96% und Spezifität von 72% berechnet. Die Gewissheit der Evidenz wurde mit *moderat* bewertet.

Cooper-Patrick et al. (1994) (101) entwickelte den *SIS-Q* und berichtete eine Sensitivität von 84%. Die berechnete Spezifität ergab einen Wert von 75%. Die Gewissheit der Evidenz wurde als *moderat* eingeschätzt.

Olfson et al. (1996) (102) prüfte das *Feeling suicidal - Suizidgedanken* Item. Die Sensitivität ergab einen Wert von 86%, die Spezifität ergab einen Wert von 99%. Die Gewissheit der Evidenz wurde als *hoch* gewertet.

4.5.2 ZIEL-ZUSTAND SUIZIDPLAN

Olfson et al. (1996) (102) untersuchte das *Feeling suicidal - Suizidplan* Item, mit einer berichteten Sensitivität von 83% und Spezifität von 98%. Des Weiteren untersuchte er den Indextest *Thoughts of death* (Sensitivität 100%, Spezifität 81%) und *Wishing you were dead* (Sensitivität 92%, Spezifität 93%) bezüglich des Ziel-Zustandes Suizidplan. Die Gewissheit der Evidenz wurde als *hoch* eingestuft.

Van Spijker et al. (2014) (106) entwickelte den *SIDAS* und berichtete eine Sensitivität von 86% und Spezifität von 67% bei einem Cut-Off-Wert von eins, sowie eine Sensitivität von 40% und Spezifität von 96% für einen Cut-Off-Wert von 21 bezüglich des Ziel-Zustandes Suizidplan. Die Gewissheit der Evidenz wurde mit *gering* bewertet.

4.5.3 ZIEL-ZUSTAND SUIZIDVERSUCH

Calear et al. (2020) (107) entwickelte den *RMTS-S* und untersuchte den Screener bezüglich der Diskrimination zwischen Personen mit und ohne Suizidversuch. Es wurde ein Cut-Off-Wert von ≥ 9 gewählt. Die Autoren berichteten für das *Development sample* eine Sensitivität von 86% und Spezifität von 83%. Für das *Validation sample* wurde eine Sensitivität von 87% und eine Spezifität von 70% beschrieben. Zudem wurde im *Development sample* die diagnostische Genauigkeit des *SBQ-R* bei einem Cut-Off-Wert von ≥ 13 untersucht und mit einer Sensitivität von 82% und einer Spezifität von 83% angegeben. Das Grade Rating ergab eine Einschätzung der Gewissheit der Evidenz als *moderat*.

Millner et al. (2015) (105) untersuchte die *Gate question suicide attempt*. Dies resultierte in einer Sensitivität von 89% und Spezifität von 94% verglichen mit einer aufwändigen Kodierung des Ziel-Zustandes Suizidversuch. Die Gewissheit der Evidenz wurde als *moderat* bewertet.

Van Spijker et al. (2014) (106) evaluierte den *SIDAS* hinsichtlich der Fähigkeit Suizidversuche zu detektieren. Bei einem Cut-Off-Wert von eins wurde eine Sensitivität von 84% und eine Spezifität von 64% berichtet. Bei einem Cut-Off-Wert von 21, wurde

eine Sensitivität von 50% und eine Spezifität von 95% berichtet. Die Gewissheit der Evidenz wurde mit *gering* bewertet.

4.5.4 ZIEL-ZUSTAND SUIZIDALITÄT

Uebelacker et al. (2011) (103) untersuchte Sensitivität und Spezifität des *Item 9* des *PHQ-9*. Die Sensitivität ergab 69% und die Spezifität 84%. Die Gewissheit der Evidenz wurde als *moderat* eingestuft.

4.6 SYNTHESE DER ERGEBNISSE

Die Maße der diagnostischen Genauigkeit Sensitivität und Spezifität werden in *Abbildung 7* (Studien mit der Population Allgemeinbevölkerung) und *Abbildung 8* (Studien mit der Population der Allgemeinmedizin) in einem Forest Plot visualisiert. Die Abbildungen wurden von MF erstellt und sind Teil der Publikation (83).

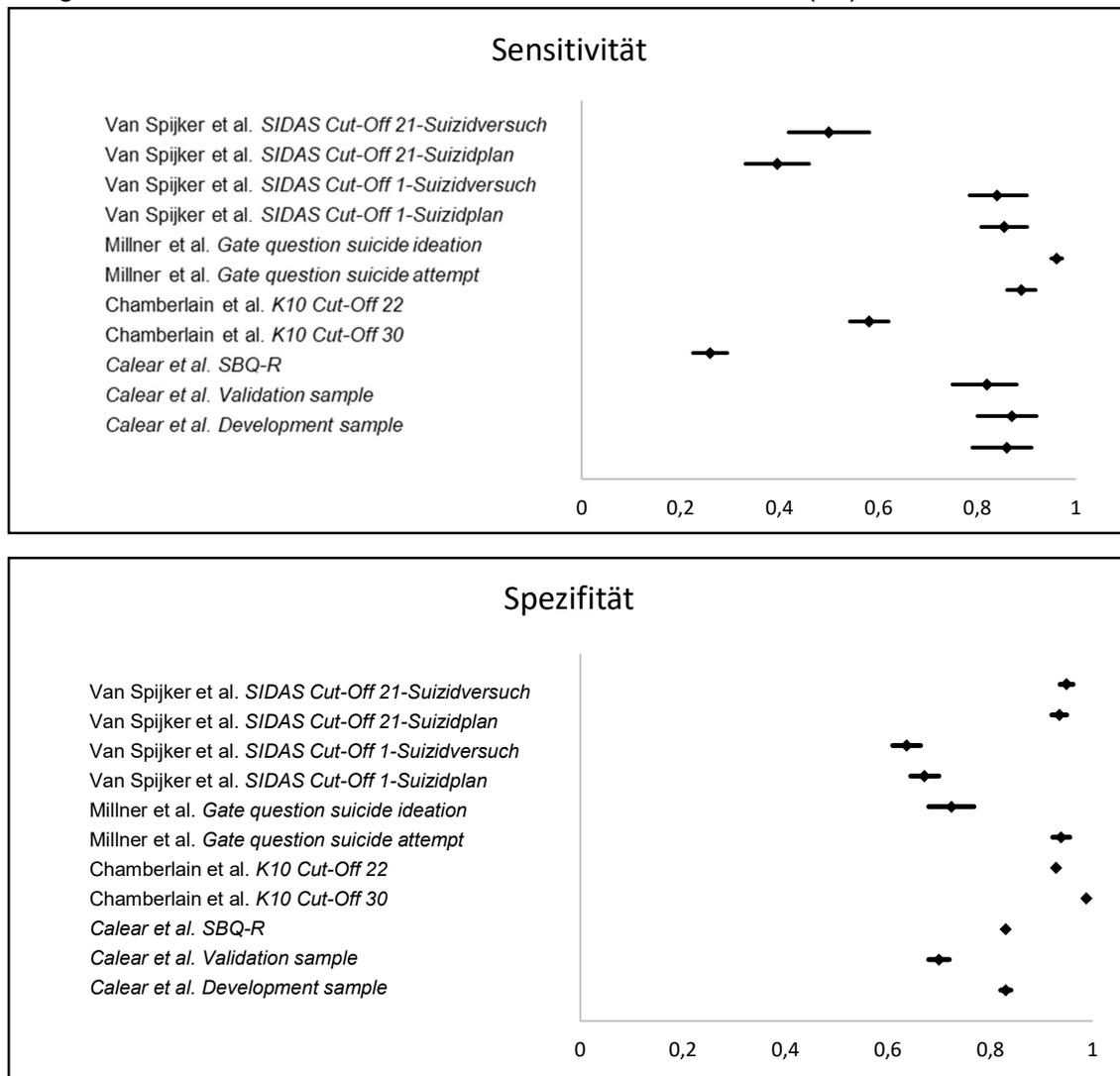


Abbildung 7: Sensitivität und Spezifität in der Allgemeinbevölkerung dargestellt in einem Forest Plot

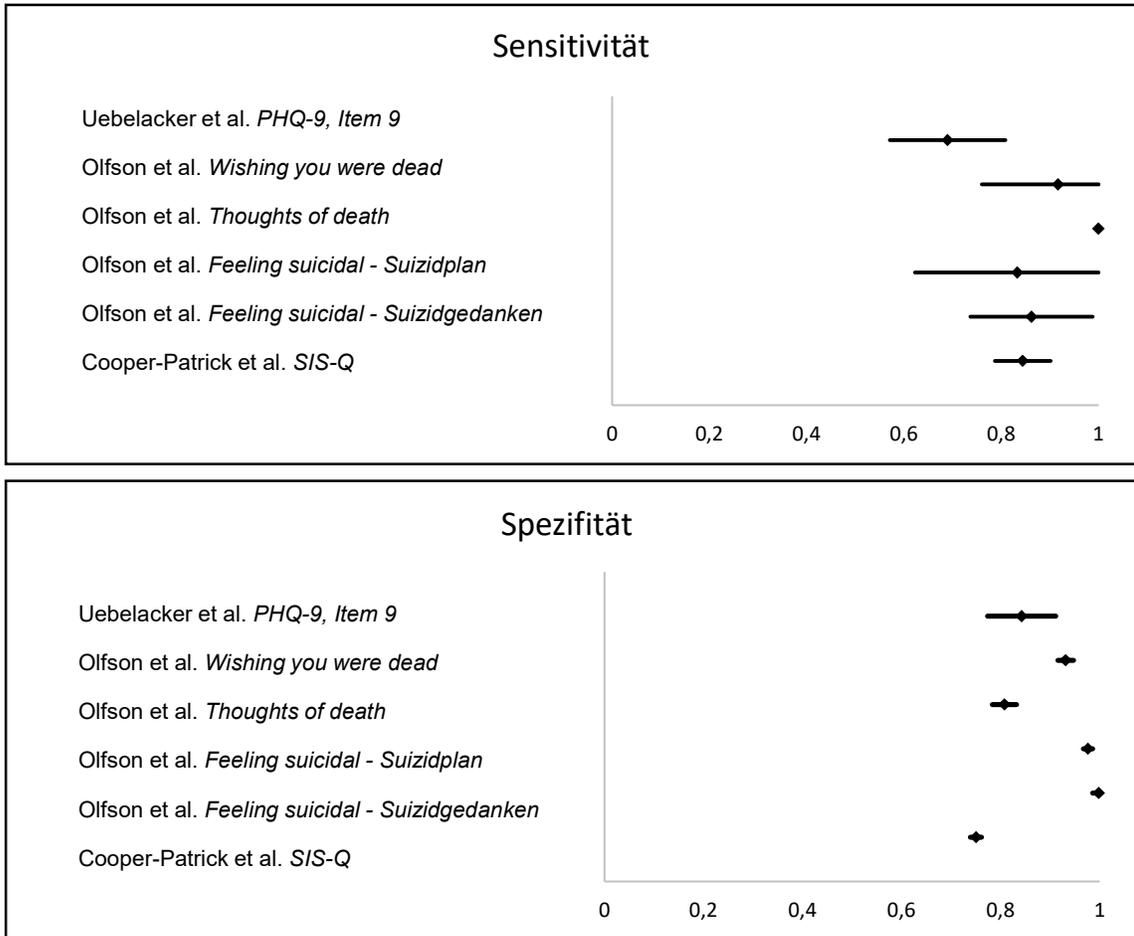


Abbildung 8: Sensitivität und Spezifität in der Allgemeinmedizin dargestellt in einem Forest Plot

5 DISKUSSION

5.1 ALLGEMEINES

„Mit Ausnahme weniger Dinge, die ganz an der Oberfläche liegen, beruht alle Diagnostik auf Wahrscheinlichkeitsrechnung“ (118).

Es wurden 15 Studien identifiziert, die kurze Screening Instrumente für Suizidalität in der Allgemeinmedizin oder Allgemeinbevölkerung untersuchten. Nur sieben dieser Studien untersuchten Screening Instrumente bezüglich ihrer diagnostischen Genauigkeit und konnten in die systematische Übersichtsarbeit eingeschlossen werden. Im Vergleich zur einer vorhergehenden Übersichtsarbeit zu diesem Thema (74) wurden in der Population der erwachsenen allgemeinmedizinischen Patienten zwei weitere Studien identifiziert (101, 103) sowie vier Studien in der Allgemeinbevölkerung (104-107). (83)

5.2 DISKUSSION DER METHODEN

5.2.1 METHODISCHE STÄRKEN

5.2.1.1 DIE SYSTEMATIK

Bei durchgeführter wissenschaftlicher Arbeit handelt es sich um eine systematische Übersichtsarbeit. Diese Methodik nimmt für sich in Anspruch, aufgrund der systematischen Literaturrecherche den aktuellen Forschungsstand zu einem Thema wieder spiegeln zu können (57).

5.2.1.2 STUDIEN ZUR DIAGNOSTISCHEN GENAUIGKEIT

Eine weitere Stärke stellt die Wahl eines engen methodischen Rahmens bezüglich des Einschusses ausschließlich von Studien zur diagnostischen Genauigkeit dar. Diese Studien bedeuten einen Standardstudientyp und liefern wertvolle Informationen für die Beurteilung der Einsatzfähigkeit eines neuen diagnostischen Tests. Studien zur

diagnostischen Genauigkeit liefern Belege dafür, wie gut ein Indextest Ziel-Zustände korrekt identifiziert. Dies unterstützt eine evidenzbasierte Entscheidungsfindung von Ärzten und Patienten. (59, 60)

5.2.1.3 DAS STUDIENPROTOKOLL

Die Vorgehensweise wurde bereits in einem Studienprotokoll (ID: CRD42019122173) (81) (siehe 3.5) festgelegt. Dieses wurde im international anerkannten Register systematischer Reviews PROSPERO (76) veröffentlicht und trug wesentlich zu einer Qualitätssicherung bei.

5.2.1.4 DIE PRISMA-DTA CHECKLISTE

Durch die stete Berücksichtigung der etablierten PRISMA-DTA Checkliste (79) (mit Seitenangaben vervollständigte PRISMA-DTA Checkliste siehe Anhang G), konnte garantiert werden, keine wichtigen Inhalte zu übersehen. Damit stellte dies einen weiteren bedeutenden Aspekt der Qualitätssicherung dar.

5.2.1.5 DIE LITERATURERECHERCHE

Die Literaturrecherche wurde systematisch durchgeführt, um möglichst alle relevanten Studien zu identifizieren. Außerdem fand eine vollständige Dokumentation der Informationsquellen (siehe 3.7.1) und Suchstrategie (vollständige Suchstrategie siehe Anhang G) statt, um eine Nachvollziehbarkeit und Replizierbarkeit zu gewährleisten.

5.2.1.6 DIE STUDIENSELEKTION

Die Studienselektion (Titel- und Abstractselektion sowie Volltextselektion) wurde jeweils von zwei Rezensenten, teils unabhängig voneinander (für genauere Informationen siehe 3.8) durchgeführt. Bei Uneinigkeit stand immer ein dritter Rezensent für eine endgültige Entscheidung zur Verfügung. Auch dies trug zu einer Qualitätssicherung bei und förderte einen wissenschaftlichen Diskurs.

5.2.1.7 DIE DATENEXTRAKTION

Für die Datenextraktion wurde eine anerkannte Checkliste – die STARD Checkliste (91) (siehe Anhang G) – für Studien zur diagnostischen Genauigkeit verwendet. Dies sicherte die Extraktion wesentlicher Items.

5.2.1.8 DAS QUADAS-2 TOOL UND GRADE RATING

Die Bewertung des Verzerrungsrisikos und die Bewertung der Evidenz wurde mit dem QUADAS-2 Tool (93) und der Software GRADEpro (97) durchgeführt. Hierbei handelt es sich um etablierte und anerkannte Instrumente.

5.2.2 METHODISCHE LIMITATIONEN

5.2.2.1 DIE SYSTEMATISCHE ÜBERSICHTSARBEIT

Wir sind uns bewusst, dass Studien aufgrund der verwendeten Suchstrategie oder des fehlerhaften Ausschlusses bei der Titel- und Abstractselektion möglicherweise nicht identifiziert wurden. Daraus folgt, dass relevante Studien zu diesem Thema übersehen worden sein könnten. Dieses Risiko wurde folgendermaßen verringert: Auf Ebene der Suchstrategie wurde diesem durch die Einhaltung der Empfehlung des Buches *Systematic reviews in health care: a practical guide* von Glasziou et al. (2001) (61) und der Einbeziehung eines erfahrenen Bibliothekars in die Suchstrategie entgegengewirkt. Auf Ebene der Titel- und Abstractselektion wurde einem fehlerhaften Ausschluss relevanter Studien durch die unabhängige Durchführung der Tätigkeit von zwei Rezensenten begegnet.

5.2.2.2 STUDIEN ZUR DIAGNOSTISCHEN GENAUIGKEIT

Aufgrund der strikten methodischen Herangehensweise, die nur Studien zur diagnostischen Genauigkeit zuließ (59, 60), konnten nur wenige Studien (n = 7) in diese systematische Übersichtsarbeit eingeschlossen werden.

Wie im PRISMA Flussdiagramm (siehe *Abbildung 2*) dargestellt wurde, mussten acht Studien mit dem Hauptgrund der fehlenden Untersuchung zur diagnostischen Genauigkeit ausgeschlossen werden. Hierdurch wurde die Anzahl der Ergebnisse mehr als halbiert. Bei den beschriebenen Studien (siehe *Tabelle 17*) handelte es sich vor allem um Studien, die in der Allgemeinbevölkerung durchgeführt wurden. In dieser Population hätten – bei einer weniger strengeren Methodik – folglich deutlich mehr Studien eingeschlossen werden können. In der Population der Allgemeinmedizin wäre nur eine weitere Studie auf diesem Weg identifiziert worden (88).

5.2.2.3 DIE EINGRENZUNG DER FORSCHUNGSFRAGE

Die strenge methodische Herangehensweise bezog sich ebenso auf eine starke Eingrenzung der Forschungsfrage anhand des PICOS-Schemas. Es wurde eine Beschränkung auf Fragebögen mit maximal zwölf Items vorgenommen. Demnach mussten bereits einige Studien auf Ebene der Titel- und Abstractselektion (in diesen Studien wurde bereits im Abstract ersichtlich, dass es sich um ein Screening Instrument mit mehr als zwölf Items handelte) und sieben weitere Studien in der Volltextselektion ausgeschlossen werden. Eine Eingrenzung auf kurze Screening Instrumente erschien jedoch für ein Screening in der Allgemeinmedizin nach wie vor sinnvoll (56, 75, 119) und wird von der Joint Commission empfohlen (119).

5.2.2.4 DIE BEWERTUNG DES VERZERRUNGSRIKOS UND DER GEWISSHEIT DER EVIDENZ

Die Bewertungen des Verzerrungsrisikos und der Bedenken bezüglich der Anwendbarkeit sind von einer nicht vermeidbaren Subjektivität geprägt. Dem wurde durch ein gewissenhaftes Befolgen des *Quadas-2-Extraktionssheets* (94) sowie durch eine unabhängige Bewertung von zwei Rezensenten entgegengewirkt. Ebenso beinhaltet das

GRADE-Rating eine gewisse Subjektivität, wurde jedoch, um diese zu verringern, von zwei Rezensenten durchgeführt.

5.2.2.5 DIE HETEROGENITÄT DER STUDIENTEILNEHMER

Es wurde sowohl das Setting der Allgemeinbevölkerung als auch das Setting der Allgemeinmedizin untersucht. Die Übertragbarkeit der Ergebnisse zwischen den Settings ist eingeschränkt, wurde jedoch aufgrund der zu erwartenden geringen Anzahl an Studien aus der Allgemeinmedizin vorgenommen. Konsequenterweise fand eine getrennte Beurteilung und Betrachtung der Ergebnisse in den wesentlichen Aspekten statt. Tabellen und Abbildungen sowie Textabschnitte wurden nach den jeweiligen Settings sortiert, um so einen besseren Überblick über die Ergebnisse der jeweiligen Populationen zu gewährleisten.

5.2.2.6 DIE NARRATIVE SYNTHESE

Eine Metaanalyse konnte aufgrund der Heterogenität der Ergebnisse nicht angewendet werden. Dies hätte eine quantitative Zusammenfassung bedeutet (57). Stattdessen konnte eine qualitative Berichterstattung mittels narrativer Synthese durchgeführt werden.

5.2.3 FAZIT

Die Methodik einer systematischen Übersichtsarbeit beansprucht den aktuellen Forschungsstand zu einem Thema widerspiegeln zu können (57). Die systematische und vollständig dokumentierte Literaturrecherche gewährleistete eine Nachvollziehbarkeit und Replizierbarkeit der Ergebnisse. Das Studienprotokoll, sowie die Verwendung unterschiedlicher anerkannter Checklisten (PRISMA-DTA Checkliste, STARD-Checkliste) und Instrumente (QUADAS-2 Tool, GRADEpro) trug wesentlich zu einer Qualitätssicherung bei. Die Durchführung wichtiger Aufgaben von zwei Rezensenten und einem dritten Rezensenten, der für eine Entscheidungsfindung zur Seite stand, reduzierte die Gefahr einer subjektiven Beurteilung und diente ebenso der Qualitätssi-

cherung. Die strenge methodische Vorgehensweise bezüglich der Auswahl von Studien zur diagnostischen Genauigkeit führte zu einer geringeren Anzahl an eingeschlossenen Studien, unterstützte jedoch eine evidenzbasierte Entscheidungsfindung. Eine Eingrenzung auf kurze Screening Instrumente führte zu einem Ausschluss von Studien, wurde jedoch im Hinblick auf die Anwendbarkeit der Ergebnisse im klinischen Alltag des Allgemeinmediziners als sinnvoll erachtet. Das Einschließen von Studienteilnehmern der Allgemeinbevölkerung schränkte die Übertragbarkeit der Ergebnisse ein, wurde jedoch wegen der zu erwartenden geringen Anzahl an Studien in der Allgemeinmedizin vorgenommen. Aufgrund der Heterogenität der Ergebnisse wurde eine narrative Synthese erstellt (99). Eine Metaanalyse, und damit eine Berechnung von gepoolten Effektschätzern (57), konnte hingegen nicht vorgenommen werden.

5.3 DISKUSSION DER ERGEBNISSE

5.3.1 EINORDNUNG DER ERGEBNISSE

5.3.1.1 SENSITIVITÄT UND SPEZIFITÄT

Die Sensitivität ist bei einem Screening Instrument besonders wichtig (103), da eine hohe Sensitivität bedeutet, dass kein Patient mit Suizidalität übersehen wird, d.h. dass die Anzahl der *falsch negativen* möglichst gering ist (120). Ein „Übersehen“ eines Patienten hingegen könnte in einem solchen Fall schwerwiegende Konsequenzen mit sich bringen (103). *Runeson et al. (2017)* legte folgende Mindestanforderung für Sensitivität und Spezifität eines Screening Instruments für Suizidalität pragmatisch fest (64): Die Sensitivität sollte über 80% liegen und die Spezifität mindestens 50% betragen. Diese Werte erreichten neun der elf Indextests: 1. *Gate question suicide ideation*, 2. *Gate question for suicide attempt*, 3. *SIS-Q*, 4. *Feeling suicidal – Suizidgedanken und Suizidplan*, 5. *Thoughts of death*, 6. *Wishing you were dead*, 7. *SIDAS – Suizidplan und Suizidversuch* mit einem Cut-Off Wert von eins, 8. *RMTS-S* und 9. *SBQ-R*.

5.3.1.2 DAS GRADE RATING

Von den Screening Instrumenten wurden zehn der elf Instrumente mit einem *moderaten* bis *hohen* Rating bewertet. Der *SIDAS* erhielt die Bewertung *gering*.

Kritisch ist hier zu bedenken, dass jeder Indextest nur in einer einzigen Studie untersucht wurde und hierdurch die Aussagekraft von GRADE eingeschränkt sein könnte.

5.3.1.3 FRAGEBÖGEN MIT NUR EINEM ITEM

Bei sechs der elf Screening Instrumente handelte es sich um Fragebögen mit nur einem Item (ausgenommen *SIDAS*, *K10*, *SIS-Q*, *RMTS-S*, *SBQ-R*). Screening Instrumente für Suizidalität mit nur einem Item werden in der Literatur kontrovers diskutiert (105, 121-123).

In der Studie von *Millner et al. (2015)* ergaben beide Indextests mit nur einem Item annehmbare Sensitivitäten, Spezifitäten, PPV und NPV, aber der Autor riet davon ab, diese Fragebögen zu verwenden, da seines Erachtens Fragebögen mit nur einem Item große Bereiche von Suizidalität nicht erfassten, und seine statistischen Simulationen zeigten, dass diese eher zu einer fehlerhafteren Klassifizierung führten als Fragebögen mit mehreren Items (105). Im Gegensatz dazu untersuchte *Simon et al. (2013)* das *Item 9 des PHQ-9* und berichtete, dass die Antwort auf *Item 9* sowohl Patienten mit einem erhöhten Risiko für Suizidversuch als auch Suizid identifizieren konnte (123). Die Anwendung des *Item 9 des PHQ-9* als Screening Instrument für Suizidalität wird bereits in einer Leitlinie empfohlen (124).

5.3.1.4 DER ZIEL-ZUSTAND

In der psychologischen Modellierung wird Suizidalität als ein kontinuierlicher Prozess beschrieben (125), der mit Suizidgedanken beginnt und bis hin zu einem Suizid fortschreiten kann. In der Allgemeinmedizin könnte eine Früherkennung sinnvoll erscheinen. Würden suizidale Patienten im Anfangsstadium identifiziert, könnte eine weitere Progression zu einem *Suizidplan* oder *Suizidversuch* mittels direkter Intervention verhindert werden (2). Auch der Ziel-Zustand *Suizidalität* enthält den Ziel-Zustand *Suizidgedanken* und könnte dementsprechend verwendet werden. Folgende Fragebögen verwendeten beschriebene Ziel-Zustände: *K10*, *Gate question suicide ideation*, *SIS-Q*, *Feeling suicidal - Suizidgedanken*, *PHQ-9 - Item 9*.

5.3.1.5 THEMATISIEREN DER SUIZIDALITÄT

Die Studien der beiden Fragebögen *K10* und *SIS-Q* stellten als Vorteil heraus, dass diese Screening Instrumente Suizidalität nicht explizit thematisieren und so den Patienten nicht direkt mit Suizidalität konfrontieren (101, 104). Die S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression empfiehlt jedoch ein direktes thematisieren von Suizidalität (37): Eine Erfassung von Suizidalität sollte bei Patienten mit Depression sowohl beim Erstkontakt, als auch im weiteren Behandlungsverlauf regelmäßig erfolgen (37); siehe *Tabelle 15*. Ein direktes Ansprechen wird auch in britischen und US-amerikanischen Leitlinien empfohlen (38, 119, 124, 126). Dabei konnten Bedenken, durch ein Screening Suizidalität auszulösen, ausgeräumt werden (127, 128). Auch *Teismann et al. (2014)* empfiehlt bei Anzeichen von Suizidalität diese direkt anzusprechen (129).

Tabelle 15: Erfassung der Risikomerkmale nach Empfehlung der S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression (37)

„Haben Sie in letzter Zeit daran denken müssen, nicht mehr leben zu wollen?“

„Häufiger?“

„Haben Sie auch daran denken müssen, ohne es zu wollen? Haben sich Suizidgedanken aufgedrängt?“

„Konnten Sie diese Gedanken beiseiteschieben?“

„Haben Sie konkrete Ideen, wie Sie es tun würden?“

„Haben Sie Vorbereitungen getroffen?“

„Umgekehrt: Gibt es etwas, was Sie davon abhält?“

„Haben Sie schon mit jemandem über Ihre Suizidgedanken gesprochen?“

„Haben Sie jemals einen Suizidversuch unternommen?“

„Hat sich in Ihrer Familie oder Ihrem Freundes- und Bekanntenkreis schon jemand das Leben genommen?“

Daher kann ein direktes Thematisieren von Suizidalität empfohlen werden.

5.3.1.6 SCHUTZFAKTOREN

Einen weiteren Aspekt stellt die Berücksichtigung von Schutzfaktoren bei der Einschätzung von Suizidalität dar. Hoffnung und Optimismus tragen zur psychischen Gesundheit bei (2, 130) und sollten in die Exploration von Suizidalität einbezogen werden (37, 88). In einer Studie konnte gezeigt werden, dass Optimismus einen Schutzfaktor für Suizidalität bei allgemeinmedizinischen Patienten darstellt (25). In dieser systematischen Übersichtsarbeit konnte kein Screening Instrument eingeschlossen werden, welches Schutzfaktoren in die Einschätzung der Suizidalität einfließen lässt. Ein Screening Instrument, welches Schutzfaktoren in die Bewertung einbezieht, ist z.B. der *P4* (88). Dieser musste im vorliegenden Review aufgrund fehlender Maße zur diagnostischen Genauigkeit im Stadium der Volltextanalyse ausgeschlossen werden.

5.3.1.7 SELBSTBEURTEILUNG ODER FREMDBEURTEILUNG

Einen weiteren Aspekt eines Fragebogens stellt dar, ob dieser selbstbeurteilt oder fremdbeurteilt verwendet wird, da dies zu unterschiedlichen Testergebnissen führen

kann (131, 132). So könnte im Allgemeinen angenommen werden, dass die Beurteilung eines Arztes eine bessere Einschätzung eines Symptoms oder einer Erkrankung, aufgrund seiner Professionalität abgeben würde. Eine Studie zeigte jedoch, dass selbstbeurteilte Suizidalität eine stärkere prädiktive Bedeutung als fremdbeurteilte habe (132), und diese beiden nur in geringem Maße korrelieren (131). Patienten schätzten ihre Suizidalität selbst öfter als schwerer ein als Ärzte (133).

5.3.1.8 SCREENING VON RISIKOPATIENTEN

Screening von Risikopatienten wird bereits von mehreren Fachgesellschaften empfohlen (37, 126, 134).

In diesem Zusammenhang sollte auch bedacht werden, dass eine Stichprobe von allgemeinmedizinischen Risikopatienten (selektioniertes Patientenkollektiv, z.B. Patienten mit Depression) und eine Stichprobe mit allen allgemeinmedizinischen Patienten (unausgelesenes Patientenkollektiv) unterschiedliche Prävalenzen bezüglich Suizidalität aufweisen würden. Denn die Prävalenz einer Erkrankung beeinflusst direkt die Werte PPV und NPV (135). PPV und NPV sind hierbei für Ärzte besonders interessant, da sie beschreiben, mit welcher Wahrscheinlichkeit, ein Patient mit einem positiven (negativen) Testergebnis wirklich erkrankt (nicht erkrankt) ist (135).

Bei den Populationen Allgemeinmedizin (unausgelesenes Patientenkollektiv) und Allgemeinbevölkerung handelt es sich normalerweise um Settings mit einer niedrigen Prävalenz von Suizidalität (16, 17, 19), z.B. 1% (102) und 5% (104). Der sich daraus ergebende NPV von 96% (104) - 100% (102) könnte die geringe Prävalenz von Suizidalität mit einer Wahrscheinlichkeit gesund zu sein von 95% (104) - 99% (102) widerspiegeln. Der zugehörige PPV Wert von ca. 6% - 52% bedeutet gleichzeitig, dass 48% - 94% der Studienteilnehmer mit einem positiven Screening Befund eine falsch positive Diagnose erhielten. Daher ist die Vorhersage, wie bei jedem anderen/r Symptom/Erkrankung im Niedrigprävalenzbereich, eingeschränkt (136).

Die Population der Risikopatienten würde eine höhere Prävalenz von Suizidalität aufweisen und dies sich auf die prädiktiven Vorhersagewerte auswirken.

Ein Rechenbeispiel in Anlehnung an *Pokorny (1983) (137)* zitiert nach *Carter et al. (2017) (138)* mit folgenden hypothetischen Annahmen soll dies verdeutlichen.

Bei der Annahme eines hypothetischen Indextests mit einer Sensitivität und Spezifität von jeweils 99% und einer Prävalenz der Erkrankung in der zu untersuchenden Population von 1% (Niedrigprävalenzbereich) ergäbe sich ein PPV von 50% und ein NPV von 98%. Das heißt, obwohl Sensitivität und Spezifität außerordentlich gut sind, ist bei dieser niedrigen Prävalenz der Erkrankung nur jeder zweite Patient mit positivem Testergebnis auch wirklich erkrankt.

Bei der Annahme eines Indextests mit selbiger Sensitivität und Spezifität von jeweils 99% aber einer erhöhten Prävalenz der Erkrankung von 35% (Prävalenz von Suizidalität bei allgemeinmedizinischen Patienten mit Depression (103)) in der zu untersuchenden Population, würde sich eine PPV von 98% und NPV von 99% ergeben.

Daher scheint auch aus statistischen Gründen ein Screening von Risikopatienten, d.h. ein Screening in einer Patientengruppe mit einer höheren Prävalenz von Suizidalität, sinnvoller, da so vor allem der PPV steigt.

Ausschließlich die Studie von *Übelacker et al. (2011) (103)* untersuchte Risikopatienten (Patienten mit Depression) hinsichtlich ihrer Suizidalität. Die Patienten wurden einem Vortest für Depression unterzogen, und nur die positiv depressiven Patienten einem Screening für Suizidalität unterzogen. Die Prävalenz von Suizidalität erreichte in dieser Patientengruppe 35%. Die Sensitivität betrug 69%, die Spezifität 84%, der PPV 70% und der NPV 83%.

In den Studien der Allgemeinbevölkerung wiesen zwei Studien mit Onlineumfragen (105, 106) höhere Prävalenzen bezüglich Suizidalität auf, als in dieser Population zu erwarten wäre. Dabei wies in der Studie von *Millner et al. (2015) (105)* die *Gate question suicide Ideation* in einer Patientengruppe mit einer Prävalenz 75% von Suizidgeanken eine Sensitivität von 96%, eine Spezifität von 72%, einen PPV von 91% und einen NPV von 86% auf. In dieser Hochprävalenzgruppe zeigte dieses Screening Instrument akzeptable Maße der diagnostischen Genauigkeit.

5.3.1.9 ÜBERBLICK ÜBER DIE EINORDNUNG DER ERGEBNISSE

Tabelle 16 gibt einen Überblick über die Einordnung der Ergebnisse der verschiedenen Screening Instrumente.

Tabelle 16: Überblick über die Einordnung der Ergebnisse

Indextext	Sensitivität $\geq 80\%$ und Spezifität $\geq 50\%$	GRADE moderat oder hoch	Mehr als ein Item	Ziel-Zustand Sui- zidgedanken oder Suizidalität	Thematisieren von Suizidalität	Schutz- faktoren	Selbstbeurteilungs- fragebogen	Risiko- patien- ten
RMTS-S	x	x	x	0	x	0	x	0
SBQ-R	x	x	x	0	x	0	x	0
K10 Cut-Off 30	0	x	x	x	0	0	0	0
K10 Cut-Off 22	0	x	x	x	0	0	0	0
Gate question suicide attempt	x	x	0	0	x	0	x	0
Gate question suicide ideation	x	x	0	x	x	0	x	0
SIDAS Cut-Off 1 – Suizidplan	x	0	x	0	x	0	x	0
SIDAS Cut-Off 1 – Suizidversuch	x	0	x	0	x	0	x	0
SIDAS Cut-Off 21 – Suizidplan	0	0	x	0	x	0	x	0
SIDAS Cut-Off 21 – Suizidversuch	0	0	x	0	x	0	x	0
SIS-Q Cut-Off 1	x	x	x	x	0	0	0	0
Feeling suicidal – Suizidgedanken	x	x	0	x	x	0	x	0
Feeling suicidal – Suizidplan	x	x	0	0	x	0	x	0
Thoughts of death	x	x	0	0	x	0	x	0
Wishing you were dead	x	x	0	0	x	0	x	0
PHQ-9, Item 9	0	x	0	x	x	0	x	x

x= trifft zu, 0 = trifft nicht zu

5.3.2 LIMITATIONEN BEZÜGLICH DER ERGEBNISSE

5.3.2.1 VERWENDETE INDEXTESTS UND REFERENZSTANDARDS

Alle in diese Übersichtsarbeit eingeschlossenen Studien bewerteten unterschiedliche Indextests. Jeder Indextest wurde in jeweils nur einer Studie beurteilt. Dies führte zu einer limitierten Aussagekraft der Ergebnisse. Zudem verwendeten die eingeschlossenen Studien verschiedene Referenzstandards. Da es bei der Beurteilung der Suizidalität keinen einheitlichen Goldstandard gibt, ist ein Vergleich der Werte der diagnostischen Genauigkeit der Indextests erschwert.

5.3.2.2 DIE GERINGE ANZAHL DER STUDIEN IN DER ALLGEMEINMEDIZIN MIT VORTEST

Des Weiteren ist zu bedenken, dass Leitlinien ein Screening von Patienten mit einem erhöhten Risiko für Suizidalität (z.B. Patienten mit Depression) empfehlen (37, 38). Dabei konnte nur eine Studie identifiziert werden (103), die in einem Vortest Patienten mit Depression detektierte, und nur diese einem Screening für Suizidalität unterzog. Die Empfehlung der Leitlinien muss in weiteren Studien untersucht werden.

5.3.2.3 DIE STUDIEN TEILNEHMER DER EINGESCHLOSSENEN STUDIEN

An den eingeschlossenen Studien nahmen insgesamt mehr Frauen als Männer teil. Dies kann zu einer Verzerrung der Ergebnisse führen. Zudem wiesen zwei Studien der Allgemeinbevölkerung mit dem Setting Onlineumfrage eine höhere Prävalenz von Suizidalität auf als in dieser Population zu erwarten wäre. Auch dies könnte zu einer Verzerrung des Ergebnisses der Screening Instrumente beitragen. Folglich wurde dies in der Interpretation der Ergebnisse berücksichtigt.

5.3.2.4 DIE ERGEBNISSE DER BEWERTUNG DES VERZERRUNGSRIKOS UND DES GRADE-RATINGS

Die Einschätzung des Verzerrungsrisikos kam zu einem moderaten Ergebnis. Das bedeutet, dass eine Wiederholung jeweiliger Studien zu anderen Ergebnissen führen könnte. Die Bewertung der Studien anhand des QUADAS-Tools wurde in der Diskussion berücksichtigt.

Ähnliche Limitationen gelten für das GRADE-Rating. Auch hier erreichten nur zwei der Studien ein hohes Vertrauen in die Ergebnisse, die anderen Studien wiesen ein geringes bis moderates Rating auf, was bedeutet, dass weitere Studien zu anderen Ergebnissen kommen könnten.

5.3.3 FAZIT

5.3.3.1 SCHLUSSFOLGERUNG FÜR DIE SCREENING INSTRUMENTE

Folgende Schlussfolgerungen können für die identifizierten Screening Instrumente aus der Einordnung der Ergebnisse gezogen werden:

5.3.3.1.1. DER RMTS-S

Der *RMTS-S* stellt einen Selbstbeurteilungsfragebogen dar, der mehrere Items verwendet. In der Studie von Caelear et al. (2020) (107) wurde das Screening Instrument entwickelt um das Suizidrisiko einzuschätzen. Die diagnostische Genauigkeit wurde jedoch nur bezüglich eines vorangegangenen Suizidversuchs ermittelt. Dabei erzielte das Screening Instrument die geforderten Werte für Sensitivität und Spezifität. Der PPV erreichte jedoch nur Werte von 11-19%. Dies bedeutet, dass ein positiv Getesteter in $\geq 80\%$ eine falsch positive Diagnose erhält. Der NPV zeigte durchgehend hohe Werte von 99% und könnte damit zum Ausschluss des Ziel-Zustandes genutzt werden. Jedoch ist dabei zu bedenken, dass dies auch auf die niedrige Prävalenz zurückzuführen sein könnte.

Insgesamt wäre eine Evaluation in einem Setting mit höherer Prävalenz von Suizidalität und die Untersuchung der diagnostischen Genauigkeit mit dem Ziel-Zustand Suizidalität interessant. Die Autoren der Studie können sich den Einsatz des *RMTS-S* als Screener für eine Triage in der Allgemeinmedizin vorstellen (107).

5.3.3.1.2. DER SBQ-R

Der *SBQ-R* von *Osman et al. (2001)* (139) stellt ein international weit verbreitetes Screening Instrument dar (65). Es handelt sich um einen Selbstbeurteilungsfragebogen der vier Items enthält. In der Studie von *Calear et al. (2020)* wurde er bezüglich Sensitivität und Spezifität untersucht. Die pragmatisch geforderten Werte von *Runeson et al. (2017)* (64) mit einer Sensitivität > 80% und Spezifität > 50% konnten erreicht werden. Die diagnostische Genauigkeit wurde jedoch nur bezüglich des Ziel-Zustandes Suizidplan ermittelt. Der PPV erreichte einen Wert von 18%. Dies bedeutet, dass mehr als acht von zehn Personen eine falsch positive Diagnose erhielten. Der NPV zeigte einen Wert von 99% und könnte damit zum Ausschluss des Ziel-Zustandes genutzt werden. Hierbei muss bedacht werden, dass dies auch auf die niedrige Prävalenz (n = 4%) zurückzuführen sein könnte.

Abschließend wäre eine Untersuchung des *SBQ-R* in einer Studie der Allgemeinmedizin wünschenswert. Die englische Version des Fragebogens ist frei verfügbar und der Einsatz als Screening Instrument wird bereits von der Joint Commission empfohlen.

5.3.3.1.3. DER K10

Der *K10* kann aufgrund der geringen Sensitivität für ein Screening von Suizidalität nicht empfohlen werden. Der NPV von 96% - 98% könnte die geringe Prävalenz von Suizidalität mit einer Wahrscheinlichkeit gesund zu sein von 95% (104) widerspiegeln. Der zugehörige PPV von 30% - 52% bedeutet gleichzeitig, dass 48% - 70% der Studienteilnehmer mit positivem Screening Befund eine falsch positive Diagnose erhielten. Des Weiteren kann dieses Screening Instrument in aktueller systematischer Übersichtsarbeit nicht für die weitere Evaluation bezüglich der Anwendung als Scree-

ning Instrument für Suizidalität in der klinischen allgemeinmedizinischen Praxis empfohlen werden, da eine einheitliche Empfehlung von Fachgesellschaften besteht, Suizidalität direkt anzusprechen (37, 38, 119, 124).

5.3.3.1.4. DIE GATE QUESTION SUICIDE IDEATION UND GATE QUESTION SUICIDE ATTEMPT

Die beiden Indextests von *Millner et al. (2015)* (105) weisen akzeptable Werte der diagnostischen Genauigkeit auf. In dieser Studie handelte es sich nicht (wie bei einem Setting in der Allgemeinbevölkerung zu erwarten wäre) um ein Niedrigprävalenzsetting (Prävalenz von $n = 37\%$ für Studienteilnehmer mit Suizidversuch und Prävalenz von $n = 75\%$ für Studienteilnehmer mit Suizidgedanken). Die Sensitivität des Indextests für Suizidgedanken erreichte einen Wert von 96%. Bei einem positiven Screening Befund betrug die Wahrscheinlichkeit für eine richtige Klassifizierung über 90% (PPV = 91%). Daher erhielten nur 9% eine falsch positive Diagnose. Auch bezüglich der Fähigkeit eine Diagnose auszuschließen, schnitt der Test mit einem Wert von NPV = 86% akzeptabel ab. Der Autor empfiehlt jedoch selbst Fragebögen mit mehreren Items (siehe 5.3.1.3 Fragebögen mit nur einem Item).

5.3.3.1.5. DER SIDAS

Da das GRADE Rating des *SIDAS* mit *gering* ausfiel, ist das Vertrauen in die Ergebnisse dieses Tests eingeschränkt. Eine Untersuchung in weiteren Studien mit allgemeinmedizinischen Risikopatienten wäre interessant, um eine Einsetzbarkeit des Screening Instruments in der allgemeinmedizinischen Praxis beurteilen zu können. Eine Evaluation mit einem strukturierten Interview als Referenzstandard wäre empfehlenswert.

5.3.3.1.6. DER SIS-Q

Der *SIS-Q* stellte den einen Fragebogen dar, der sowohl die pragmatisch gewählten Mindestwerte von *Runeson et al. (2017)* (64), als auch ein moderates GRADE Rating sowie mehrere Items enthielt. Der NPV des *SIS-Q* betrug 99%. Dieser hohe Wert

könnte jedoch die geringe Prävalenz von Suizidgedanken in diesem Setting widerspiegeln – mit einer Wahrscheinlichkeit gesund zu sein von 97%. Mit einem PPV von 8% erzielte der *SIS-Q* einen geringen Wert, d.h. 92% der positiven Screening Befunde erwiesen sich als *falsch positiv*. Da eine einheitliche Empfehlung von Fachgesellschaften besteht, Suizidalität direkt anzusprechen, kann dieses Screening Instrument in aktueller systematischer Übersichtsarbeit nicht für die weitere Evaluation hinsichtlich der Anwendung als Screening Instrument für Suizidalität in der klinischen allgemeinmedizinischen Praxis empfohlen werden (37, 38, 119, 124).

5.3.3.1.7. FEELING SUICIDAL, THOUGHTS OF DEATH UND WISHING YOU WERE DEAD

Die Indextests von *Olfson et al. (1996)* (102) wiesen unterschiedliche Werte der diagnostischen Genauigkeit auf. Der Indextest *Feeling suicidal - Suizidgedanken* schnitt hinsichtlich Sensitivität, Spezifität, PPV und NPV akzeptabel ab. Zu bedenken ist, dass es sich um Indextests mit nur jeweils einem Item handelt und dass die Studie in einem Niedrigprävalenzsetting mit einer Prävalenz von Suizidalität von 1% - 3% durchgeführt wurde. Dies könnte sich im durchweg hohen NPV von 100% und einem niedrigen PPV (6%, 14%, 30%) widerspiegeln. Der Indextest *Feeling suicidal - Suizidgedanken* stellt hierbei mit einem PPV von 76% eine Ausnahme dar.

5.3.3.1.8. DER PHQ-9 ITEM 9

Auch das *Item 9* des *PHQ-9* stellt einen Fragebogen mit nur einem Item dar. Jedoch ist hier herauszustellen, dass der *PHQ-9* bereits als Screening Instrument für Depression in den Praxisalltag integriert ist (103). Der Indextest wies, mit Ausnahme der Sensitivität (69%), durchgehend eine akzeptable diagnostische Genauigkeit auf.

Die Sensitivität von 69% bedeutet, dass dieses Screening Instrument nur zwei Drittel der Patienten mit Suizidalität identifizieren würde. Ein Arzt müsste demnach auch bei einem negativen Screening Befund, aber weiterhin bestehendem Verdacht oder vorhandenen Risikofaktoren eine weitere Abklärung vornehmen (103).

Ein PPV von 70% würde dabei bedeuten, dass sieben von zehn Patienten mit einem positiven Screening Ergebnis wirklich suizidal wären. Mit einem NPV von 83%, wäre die Wahrscheinlichkeit bei einem negativen Screening Befund 83%, dass ein Patient *richtig negativ* eingestuft wurde und damit nicht suizidal ist. D.h. ungefähr acht von zehn Patienten würden richtig klassifiziert werden.

Zu bedenken ist, dass sowohl die jeweiligen Werte von Sensitivität und PPV (zum Einschluss einer Diagnose), als auch die Werte von Spezifität und NPV (zum Ausschluss einer Diagnose) nicht zu einem Erfassen des jeweiligen vollständigen Patientenkollektives führten. Jedoch könnte dies eventuell einen ersten Schritt darstellen.

Des Weiteren ist zu berichten, dass bereits eine prospektive Studie in der Allgemeinmedizin und psychiatrischen Versorgung zeigen konnte, dass die Antwort auf *Item 9* einen starken prädiktiven Wert für Suizidversuche und Suizid hat (123). Auch in einer US-amerikanischen Leitlinie wird das *Item 9* des *PHQ-9* bereits als Screening Instrument für Suizidalität empfohlen (124).

5.3.3.2 WEITERE SCHLUSSFOLGERUNGEN

Der NPV der Screening Instrumente lag bei über 80%. In den Studien mit einer niedrigen Prävalenz von Suizidalität könnte das die hohe Wahrscheinlichkeit gesund zu sein widerspiegeln. Jedoch waren diese Werte auch in den Studienpopulationen mit einer höheren Prävalenz von Suizidalität zu beobachten (103, 105). Würden weitere Studien diese Ergebnisse in einem Setting mit Risikopatienten bestätigen, könnten die Fragebögen dem Allgemeinmediziner helfen, Suizidalität auszuschließen. Laut dem *Curriculum für Allgemeinmedizin* stellt der Ausschluss einer Diagnose eine wichtige Aufgabe des Allgemeinmediziners dar (55). *Red flags* müssen ausgeschlossen werden, um anschließend aufmerksam abzuwarten (*watchful waiting*) (55).

Des Weiteren spielt die Wahl des Referenzstandards eine entscheidende Rolle, denn die Ergebnisse des Indextests hängen direkt von der Bewertung des Referenzstandards ab. Ein strukturiertes Interview, wie z.B. das *SCID-I* (115), stellt den höchsten Standard dar. In der Literatur werden jedoch verschiedene Screening Instrumente als

Goldstandard bezeichnet, darunter kurze Fragebögen wie der C-SSRS (140-142). Eine Einigung auf einen Goldstandard wäre ein wichtiger Meilenstein in der zukünftigen Forschung. Dies würde die Vergleichbarkeit der Ergebnisse verbessern.

Weitere vielversprechende kurze Screening Instrumente wurden bereits in der Allgemeinmedizin und Allgemeinbevölkerung evaluiert, aber in folgenden Studien nicht bezüglich ihrer diagnostischen Genauigkeit untersucht. Diese mussten in diesem Review auf Ebene des Volltextscreenings ausgeschlossen werden (56, 88, 141, 143-147). Eine Untersuchung in weiteren Studien wäre lohnend. Einen Überblick dieser Screening Instrumente gibt *Tabelle 17* (Die Tabelle wurde von MF erstellt und ist Teil der Publikation (83)). *Der SBQ-R* wurde von *Calear et al. (2020)* (107) bezüglich der diagnostischen Genauigkeit untersucht und konnte in diese systematische Übersichtsarbeit eingeschlossen werden.

Tabelle 17: Studien ausgeschlossen aufgrund fehlender Maße zur diagnostischen Genauigkeit

Autor; Jahr; Land	Test	Anzahl der Items	Population	Ziel der Studie
Batterham, P. J.; 2015; Australien (143)	Interpersonal- Psychological Theory of Sui- cidal Behaviour	12	Allgemeinbevöl- kerung	Evaluation der Inter- personal-Psychologi- cal Theory of Suicidal Behaviour
Dube, P.; 2010; USA (88)	P4	4	Allgemeinmedi- zin	Evaluation des P4
Fountoulakis, K.; 2012; Griechen- land (145)	RASS	12	Allgemeinbevöl- kerung	Entwicklung des RASS
Glaesmer, H.; 2018; Deutschland (147)	SBQ-R	4	Allgemeinbevöl- kerung	Evaluation der psychometrischen Ei- genschaften des SBQ-R
Harris, K. M.; Study 2 and 3; 2015; Australien (141)	SABCS	6	Allgemeinbevöl- kerung	Entwicklung und Vali- dierung des SABCS
Kliem, S.; 2017; Deutschland (56)	BSS-Screen	5	Allgemeinbevöl- kerung	Evaluation der psychometrischen Ei- genschaften des BSS -Screen
Turnell, A. I.; 2019; Australien (144)	Self-Hate Scale	7	Allgemeinbevöl- kerung	Entwicklung und Vali- dierung des Self-Hate Scale
Rimkeviciene, J.; Study 2; 2017; Australien, Singa- pur (146)	ACSS	5, 7, 8, (20)	Allgemeinbevöl- kerung	Evaluation der Validi- tät und Faktorenstruk- tur des ACSS

*RASS: Risk Assessment Suicidality Scale, SBQ-R: Suicide Behaviours Questionnaire - Revised, SABCS: Suicidal Affect-
Behaviour-Cognition Scale, BSS-Screen: German Beck Scale for Suicide Ideation - Screen, ACSS: Acquired Capability
for Suicide Scale*

5.4 DIE BEDEUTUNG FÜR DIE ALLGEMEINMEDIZIN

5.4.1 HERAUSSTELLUNGSMERKMALE DES FACHARZTES FÜR ALLGEMEINMEDIZIN

Um die Einordnung der Ergebnisse und die Bedeutung für die Allgemeinmedizin vorzunehmen, ist es wichtig, einige Besonderheiten des Allgemeinmediziners im Gegensatz zu anderen Fachärzten darzustellen.

Dies stellt zum einen das unausgelesene Patientenkollektiv des Allgemeinmediziners dar (148). Der Allgemeinmediziner ist eine erste Anlaufstelle für jedwede körperliche, psychische, akute oder chronische Erkrankungen oder Symptome eines Patienten (148) und stellt eine qualitativ hochstehende Grundversorgung sicher (55): Ein Patient stellt sich mit einem gewissen Anliegen (Beratungsanlass) vor. Dies resultiert ärztlicherseits in einem Beratungsergebnis und führt zu weiteren Maßnahmen (Behandlungsepisode). Dabei kann es sich hierbei um erforderliche Handlungsoptionen aus Prävention, Diagnostik, Therapie, Rehabilitation oder eine Koordination und Weitervermittlung handeln. Folgende Punkte sollten hierbei stets berücksichtigt werden (55):

- *„Gibt es Hinweise für eine akut bedrohliche Erkrankung („red flag“), die einer sofortigen Handlung bedarf und welche Differentialdiagnosen kommen in Frage?“*
- *„Welches Procedere (anschließende Diagnostik und Therapie) ist einzuleiten (oder bewusst nicht einzuleiten)?“*

Ein weiteres Herausstellungsmerkmal ist die oft langjährige Arzt-Patienten-Beziehung (148). Oft betreut der Hausarzt das soziale Umfeld des Patienten und kennt auch das häusliche Umfeld des Patienten (Hausbesuch) (148). Dies gibt dem Allgemeinmediziner die Möglichkeit, die Beschwerden seines Patienten in ein biopsychosoziales Krankheitskonzept einzuordnen (ganzheitliches Fallverständnis) (148).

Eine weitere Aufgabe des Allgemeinmediziners ist eine adäquate Versorgung anzustreben, daher eine Fehlversorgung, Unterversorgung oder Überversorgung zu vermeiden. In diesem Zusammenhang stellt die Stufendiagnostik sowie ein individuelles

Abwägen von Vor- und Nachteilen der jeweiligen Behandlungsepisode in einem ganzheitlichen Fallverständnis eine wichtige Maßnahme dar (55).

5.4.2 DIE KLINISCHE RELEVANZ DER ERGEBNISSE

Die Diagnostik von Suizidalität stellt eine extrem wichtige, aber auch extrem schwierige Aufgabe dar (148). Screening von Suizidalität ist ein erster und wichtiger Schritt (5). Für das Thema Suizidalität ist zwar der Facharzt für Psychiatrie der Spezialist, jedoch werden viele suizidale Patienten außerhalb eines psychiatrischen Settings behandelt und könnten daher von einer Versorgung in der Allgemeinmedizin profitieren (149). Fast die Hälfte der Menschen, die durch einen Suizid versterben, wurden im Monat davor bei einem Hausarzt vorstellig (47, 48).

5.4.2.1 SCREENING VON SUIZIDALITÄT IN DER ALLGEMEINMEDIZIN

Für psychische Erkrankungen ist der Allgemeinmediziner oft der erste Ansprechpartner (148). Eine wichtige Kompetenz des Allgemeinmediziners ist, sogenannte „red flags“ zu erkennen. Dies kommt bei einem Screening von Suizidalität zum Tragen. Der Allgemeinmediziner stellt die psychosomatische Grundversorgung des Patienten sicher. Dies beinhaltet Basisdiagnostik, Basistherapie und die Zusammenarbeit und Überweisung mit/an Psychotherapeuten oder einen spezialisierten Facharzt (54). Für eine optimale Versorgungsqualität ist die Kooperation von Grundversorgung (Facharzt für Allgemeinmedizin) und spezialisierter Versorgung (Facharzt für Psychiatrie) essentiell (54). Aus dieser Perspektive erscheint demnach ein Screening von Suizidalität eines selektionierten Patientenkollektivs (z.B. Patienten mit Depression) in der Allgemeinmedizin sinnvoll. Dies wird bereits von mehreren Fachgesellschaften empfohlen (37, 38, 126).

5.4.2.2 DIE AUSWAHL DES PATIENTENKOLLEKTIVS

Ein Screening von Suizidalität kann einen möglichen Nutzen und möglichen Schaden bedeuten. Anders als bei der Durchführung einer Therapie, ist bei einem Screening ein Großteil der Patienten nicht von der Erkrankung betroffen. Einen Nutzen haben folglich die Patienten, die als *richtig positiv* identifiziert werden. Ein Schaden kann durch eine weitere Abklärung bei einem *falsch positiven* Ergebnis entstehen, da dies für den Patienten eine Belastung darstellen kann. (120) Wichtig ist hier jedoch herauszustellen, dass nach aktuellem Forschungsstand durch ein Screening von Suizidalität keine Suizidalität hervorgerufen wird (127, 128).

Daher stellt sich die Frage, welche Patienten von einem Screening von Suizidalität in der Allgemeinmedizin profitieren könnten. Hier gibt es von verschiedenen Fachgesellschaften unterschiedliche Empfehlungen. Die United States Preventive Services Task Force (USPSTF) gibt aufgrund der geringen zur Verfügung stehenden Daten keine Empfehlung ab (150). Die Joint Commission empfiehlt hingegen alle Patienten in der Allgemeinmedizin hinsichtlich Suizidalität zu befragen (119). Die britischen und deutschen Leitlinien empfehlen Risikopatienten (Patienten mit Depression) bezüglich Suizidalität zu screenen (37, 38).

Die Empfehlung der Joint Commission alle Patienten in der Allgemeinmedizin zu screenen, beruht darauf, dass auch Patienten, die keine Risikopatienten für Suizidalität darstellen, Suizid begehen. Wenn nur Patienten mit Risikofaktoren beachtet werden würden, könnten Patienten ohne Risikofaktoren mit erhöhter Suizidalität übersehen werden. Dennoch weist ein Großteil der Suizidenten Risikofaktoren auf. Die Joint Commission schlägt folgendes Vorgehen vor: Der Arzt sollte in einem ersten Schritt die Risikofaktoren, die mit einem höheren Risiko für Suizidalität – wie z.B. Depression, vorangehende Suizidversuche oder soziale Isolation – einhergehen, in der Krankengeschichte eruieren. Zudem sollten alle Patienten im Wartezimmer ein Screening Instrument für Suizidalität erhalten. Dabei empfiehlt die Joint Commission kurze Screening Instrumente wie den *PHQ-9* oder den *SBQ-R*. Der Vorteil eines Screening Instruments für Suizidalität bestünde vor allem darin, das Thema Suizidalität direkt anzusprechen. Ärzte hingegen sprächen das Thema oft nur vage an. Im Anschluss kann je

nach Ergebnis des Screenings und der Risikofaktoren ein vertiefendes Screening Instrument, wie z.B. der C-SSRS eingesetzt werden, Angehörige miteinbezogen werden sowie ein spezialisierter Facharzt hinzugezogen werden. (119)

In den Studien von *Olfson et al. (1996)* und *Cooper-Patrick et al. (1994)* erhielten alle Patienten (unausgelesenes Patientenkollektiv) ein Screening Instrument für Suizidalität. In dieser Patientengruppe wies die Prävalenz von Suizidalität erwartungsgemäß einen geringen Wert auf (1% - 3%). Statistisch ist die Vorhersage, wie bei jedem anderen/r Symptom/Erkrankung im Niedrigprävalenzbereich hierbei jedoch eingeschränkt (136) (siehe 5.3.1.8).

Die Empfehlung Risikopatienten (selektioniertes Patientenkollektiv) in der Allgemeinmedizin zu screenen bezieht sich bisher vor allem auf Patienten mit Depression (37, 38, 126). Der Allgemeinmediziner spielt in der Diagnose und Therapie der Depression eine wichtige Rolle (151). Er stellt die erste Anlaufstelle dar und kennt den Patienten und sein soziales Umfeld oft seit Jahren (148). Dementsprechend kann er Veränderungen besser detektieren (151). Dabei stellt eine Depression mit einer Lebenszeitprävalenz von 20% eine relativ häufige Erkrankung dar und ist ein wichtiger Risikofaktor für Suizidalität. Knapp ein Fünftel der Patienten, die häufiger an dieser Erkrankung leiden, verstirbt an einem Suizid (151).

Die Empfehlung Risikopatienten in der Allgemeinmedizin zu screenen wurde nur in einer identifizierten Studie (103) umgesetzt. Patienten, bei denen mittels Vortest eine Depression detektiert wurde, wurden mit dem *Item 9* des *PHQ-9* bezüglich Suizidalität untersucht. Die Prävalenz in diesem selektionierten Patientenkollektiv erreichte einen Wert von 35% und wies, mit Ausnahme der Sensitivität (69%), durchgehend eine akzeptable diagnostische Genauigkeit auf.

Ein Screening von Risikopatienten (Patienten mit Depression) in der Allgemeinmedizin zeigte sich, wenn auch nur in einer Studie, als ein erster Schritt und sollte weiter – auch hinsichtlich anderer Risikogruppen – untersucht werden.

5.4.3 FACHSPEZIFISCHE LIMITATIONEN VON SCREENING FÜR SUIZIDALITÄT IN DER ALLGEMEINMEDIZIN

Eine wichtige fachspezifische Limitation stellt das für die Allgemeinmedizin typische unausgelesene Patientenkollektiv (148) dar. Die Prävalenz von Suizidalität befindet sich hier in einem Niedrigprävalenzbereich. In solch einem Kollektiv (niedrige Vortestwahrscheinlichkeit) ist die Vorhersage eines/r Symptoms/Erkrankung wie bei jedem anderen/r Symptom/Erkrankung im Niedrigprävalenzbereich eingeschränkt (136). Die Prävalenz kann jedoch durch einen Vortest (z.B. Screening für Depression) und ein damit selektiertes Patientenkollektiv angehoben werden (103).

Diese Übersichtsarbeit bezieht sich auf die Detektion von suizidalen Patienten in der Allgemeinmedizin, daher können zu anderen Settings keine Aussagen getroffen werden. Patienten, die nicht in der Allgemeinmedizin vorstellig werden, können mit dieser Methode nicht identifiziert werden und müssen mit weiteren Suizidpräventionsansätzen erfasst werden. Ein ausschließlicher Einsatz einer einzelnen Suizidpräventionsstrategie, wie der Einsatz von Screening Instrumenten in der Allgemeinmedizin, kann aufgrund der multifaktoriellen Ätiologie und Genese nicht zielführend sein (1, 36). Im Hinblick auf die Komplexität des Themenfeldes sind breitgefächerte Präventionsansätze erforderlich (1, 36).

5.4.4 FAZIT UND FORSCHUNGSAUFGABEN

Der aktuelle Forschungsstand zeigt, dass Suizidalität in der Allgemeinmedizin oft nicht erkannt wurde (49, 50). Ein Screening für Suizidalität in der Allgemeinmedizin bei einem selektierten Patientenkollektiv könnte sich als sinnvoll erweisen. Ein sensitives und spezifisches Screening Instrument könnte dem Allgemeinmediziner mehr Sicherheit in der Diagnostik geben (4) und bei der Einschätzung des weiteren Vorgehens behilflich sein (149).

Die Anzahl der eingeschlossenen Studien war gering und hätte auch bei einer weniger strengeren Methodik nur eine weitere Studie in der Population der Allgemeinmedizin ergeben. Weiter verschärft zeigt sich diese Situation hinsichtlich eines Screenings von

Risikopatienten. In diesem Setting konnte nur eine Studie – die Patienten mit Depression untersuchte – identifiziert werden. Hierbei wird ein Screening von Risikopatienten von unterschiedlichen Fachgesellschaften empfohlen (37, 38, 126).

In der Allgemeinmedizin und besonders in einem selektionierten Setting, besteht demnach weiterhin ein großer Forschungsbedarf hinsichtlich der Evaluation von kurzen Screening Instrumenten für Suizidalität.

Interessant könnte sich auch die Untersuchung anderer selektionierter Patientenkollektive (z.B. Patienten mit Angststörungen (53)) erweisen.

Aufgrund der großen Verantwortung kann mit aktueller Datenlage noch keine Empfehlung eines Screening Instruments für die Anwendung in der klinischen Praxis ausgesprochen werden. Kurze Screening Instrumente für Suizidalität müssen ihre diagnostische Genauigkeit in weiteren Studien unter Beweis stellen, um dem Allgemeinmediziner ein Instrument an die Hand geben zu können. Eine Sicherheit bezüglich des Ein- als auch des Ausschlusses von Suizidalität würde für den Allgemeinmediziner einen klinischen Nutzen bedeuten.

Ein Screening Instrument kann die Diagnose eines Arztes nicht ersetzen. Ebenso sollte ein negatives Ergebnis eines Screening Instruments bei Vorliegen von anderen Risikofaktoren für Suizidalität weiter untersucht werden (103). Das klinische Urteil des behandelnden Arztes stellt die endgültige Entscheidung dar (129).

5.4.5 AUSBLICK

Folgende strukturierte Vorgehensweise wurde für eine Etablierung eines Screening Instruments für Suizidalität in der Allgemeinmedizin angewendet (siehe *Abbildung 9*). In einem ersten Schritt wird ein selektioniertes Patientenkollektiv mit einer erhöhten Wahrscheinlichkeit für eine höhere Prävalenz von Suizidalität durch die Anwendung des *PHQ-2* gebildet (Patientenkollektiv mit möglicher Depression). Bei einem positiven Screening wird der *PHQ-9*, inklusive *Item 9* vertiefend eingesetzt (Screening der Patienten mit möglicher Depression hinsichtlich möglicher Suizidalität). Resultiert dies in

einem positiven Ergebnis, wird weiter vertiefend der C-SSRS Fragebogen eingesetzt (vertiefende Beurteilung der Suizidalität). (149) Dies entspricht der für den Allgemeinmediziner typischen Stufendiagnostik und verringert die Nachteile einer Unter- oder Überversorgung (55).

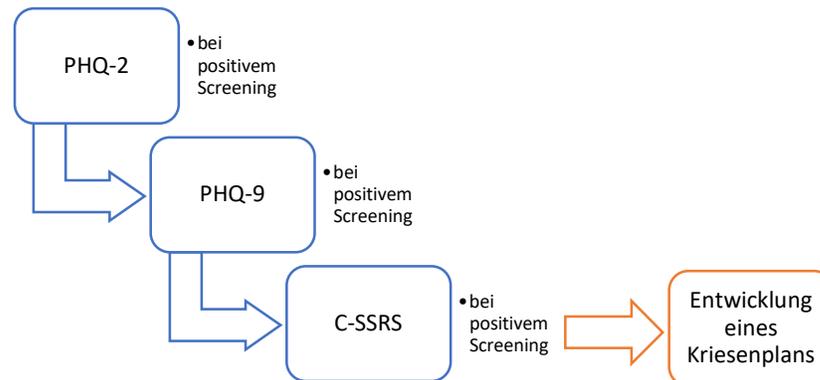


Abbildung 9: Prozess eines Screenings für Suizidalität in der Allgemeinmedizin des Kaiser Permanente Washington Primary Care System (149)

Eine Stufendiagnostik für ein Screening von Suizidalität wird ebenso von der Joint Commission empfohlen. Diese empfiehlt zunächst den Einsatz eines Screening Instruments, wie z.B. des PHQ-9 oder SBQ-R und für eine weitere Abklärung die Möglichkeit ein vertiefendes Screening, z.B. mit dem C-SSRS oder *Suicide Assessment Five-step Evaluation and Triage (SAFE-T)* (152) anzuwenden (119).

Eine weitere Forschung hinsichtlich beschriebener Stufendiagnostik wäre wünschenswert.

6 CONCLUSIO

6.1 SCHLUSSFOLGERUNG FÜR DIE FORSCHUNG

Es wurden bereits einige vielversprechende kurze Indextests zum Screenen von Suizidalität in der Allgemeinmedizin oder Allgemeinbevölkerung untersucht. Jedoch konnten nur wenige Studien identifiziert werden, die die diagnostische Genauigkeit eines Indextests prüften. Identifizierte Indextests wiesen unterschiedliche Eigenschaften (z.B. direktes Thematisieren von Suizidalität oder nicht, sowie unterschiedliche Arten der Durchführung des Tests, z.B. selbstbeurteilt oder fremdbeurteilt) auf und müssen in weiteren Studien ihre diagnostische Genauigkeit unter Beweis stellen.

Da bereits viele kurze Screening Instrumente zur Verfügung stehen, wäre eine weitere Untersuchung dieser Fragebögen sowohl hinsichtlich ihrer diagnostischen Genauigkeit, als auch hinsichtlich anderer psychometrischer Eigenschaften sinnvoll.

6.2 SCHLUSSFOLGERUNG FÜR DIE ALLGEMEINMEDIZINISCHE PRAXIS

Alle Indextests wurden in nur jeweils einer Studie untersucht, wodurch die Aussagekraft der Ergebnisse limitiert ist. Unterschiedliche kurze Screening Instrumente wurden identifiziert und zeigten eine große Bandbreite bezüglich Sensitivität, Spezifität, PPV und NPV.

Der *PHQ-9* stellt ein wichtiges Screening Instrument für die Allgemeinmedizin dar, da dieser bereits erfolgreich für Screening von Depression in den Praxisalltag integriert ist. In dieser Übersichtsarbeit konnte eine Studie eingeschlossen werden, die das *Item 9* des *PHQ-9* als Screening Instrument für Suizidalität untersuchte. Der Fragebogen sollte in die weitere allgemeinmedizinische Forschung einbezogen werden. Zum einen für ein Replizieren der diagnostischen Genauigkeit, zum anderen um weitere psychometrische Eigenschaften zu untersuchen. Eine Eingliederung in eine allgemeinmedizi-

nische Stufendiagnostik wäre interessant. Das *Item 9* des *PHQ-9* wird bereits von einer US-amerikanischen Fachgesellschaft für den Einsatz als Screening Instrument für Suizidalität empfohlen (124).

Abschließend ist festzuhalten, dass ein Screening Instrument die Diagnose eines Arztes nicht ersetzen kann. Das klinische Urteil des behandelnden Arztes stellt die endgültige Entscheidung dar (129). Dies beinhaltet die Entscheidung über eine weitere nötige Behandlung oder die Überweisung an einen psychiatrischen Facharzt (153, 154).

G ANHANG

Suchstrategie

THESAURUSSUCHE

MEDLINE

1. SUICIDE/pc [Prevention & Control], 2. Suicidal Ideation/, 3. Suicide, Attempted/,
4. 1 or 2 or 3
5. Risk Assessment/, 6. "Diagnostic Techniques and Procedures"/, 7. "Surveys and Questionnaires"/
8. Diagnostic Self Evaluation/, 9. Self Report/, 10. Psychiatric Status Rating Scales/, 11. Interviews as Topic/, 12. Mass Screening/, 13. Interview, Psychological/, 15. "Sensitivity and Specificity"/, 16. ROC Curve/, 17. "reproducibility of results"/, 18. diagnostic test/, 19. predictive value/,
14. 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13,
20. 15 or 16 or 17 or 18 or 19,
21. 4 and 14 and 20

EMBASE

1. suicide/di, pc [Diagnosis, Prevention], 2. questionnaire/, 3. interview/, 4. risk assessment/, 5. self report/, 6. diagnosis/, 7. psychological rating scale/, 8. psychologic assessment/, 9. psychological interview/, 10. diagnostic accuracy/, 11. diagnostic test/, 12. "sensitivity and specificity"/, 13. predictive value/, 14. diagnostic test accuracy study/, 15. receiver operating characteristic/, 16. suicidal behavior/, 17. suicidal ideation/, 18. suicide attempt/, 21. validity/, 22. reliability/,
19. 1 or 16 or 17 or 18, 20. 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9,
23. 10 or 11 or 12 or 13 or 14 or 15 or 21 or 22,
24. 19 and 20 and 23

COCHRANE LIBRARY

- #56 MeSH descriptor: [Suicide] this term only and with qualifier(s): [prevention & control - PC]
#57 MeSH descriptor: [Suicidal Ideation] this term only
#58 MeSH descriptor: [Suicide, Attempted] this term only
#59 #56 or #57 or #58
#60 MeSH descriptor: [Sensitivity and Specificity] this term only
#61 MeSH descriptor: [ROC Curve] this term only
#62 MeSH descriptor: [Reproducibility of Results] this term only
#63 MeSH descriptor: [Diagnostic Tests, Routine] this term only
#64 MeSH descriptor: [Predictive Value of Tests] this term only
#65 #60 or #61 or #62 or #63 or #64
#66 MeSH descriptor: [Risk Assessment] this term only
#67 MeSH descriptor: [Diagnostic Techniques and Procedures] this term only
#68 MeSH descriptor: [Surveys and Questionnaires] this term only
#69 MeSH descriptor: [Diagnostic Self Evaluation] this term only
#70 MeSH descriptor: [Self Report] this term only
#71 MeSH descriptor: [Psychiatric Status Rating Scales] this term only
#72 MeSH descriptor: [Interviews as Topic] this term only
#73 MeSH descriptor: [Interview, Psychological] this term only
#74 MeSH descriptor: [Mass Screening] this term only
#75 #66 or #67 or #68 or #69 or #70 or #71 or #72 or #73 or #74

#76 #59 and #65 and #75

PSYINDEX UND PSYCHINFO

S1: DE "Test Reliability" OR DE "Test Sensitivity" OR DE "Test Specificity" OR DE "Test Validity"
S2: DE "Suicidology" OR DE "Attempted Suicide" OR DE "Suicidal Ideation" OR DE
S3: DE "Questionnaires" OR DE "Interviews" OR DE "Risk Assessment" OR DE "Diagnosis" OR DE
"Diagnostic Interview Schedule" OR DE "Psychodiagnostic Interview" OR DE "Self- Examination (Medi-
cal)" OR DE "Screening" OR DE "Screening Tests"
(S1 AND S2 AND S3)

STICHWORTSUCHE

MEDLINE UND EMBASE

1. valid*.ti,ab., 2. reliab*.ti,ab., 3. predictive value*.ti,ab., 4. ROC curve.ti,ab., 5. sensitiv*.ti,ab., 6. speci-
ficity*.ti,ab., 7. suicide*.ti,ab., 8. self killing*.ti,ab., 9. suicidal poisoning.ti,ab., 10. self inflicted
death.ti,ab., 11. suicidal behavio?r*.ti,ab., 12. suicid* risk*.ti,ab., 13. suicid* ideation*.ti,ab., 14. suicide
attempt*.ti,ab., 15. attempted suicide*.ti,ab., 16. parasuicide*.ti,ab., 17. self harm*.ti,ab., 18. screen*
trial*.ti,ab., 19. screen*.ti,ab., 20. questionnaire*.ti,ab., 21. interview*.ti,ab., 22. risk assessment*.ti,ab.,
23. risk analysis*.ti,ab., 24. risk evaluation*.ti,ab., 25. safety assessment*.ti,ab., 26. self report*.ti,ab.,,
27. diagnos*.ti,ab., 28. diagnos* screen*.ti,ab., 29. suicidal ideation screen*.ti,ab., 30. paykel suicide
item*.ti,ab., 31. Suicide Behaviors Questionnaire.ti,ab., 32. Suicide Ideation Scale.ti,ab., 33. Suicide
Ideation Screening Questionnaire.ti,ab., 34. Beck Depression Inventory.ti,ab., 35. Hamilton Rating
Scale for Depression.ti,ab., 36. Beck fast Scan.ti,ab., 37. Suicide risk Screen.ti,ab., 38. Columbia-Sui-
cide Severity Rating Scale.ti,ab., 39. (suicide behavio?rs questionnaire and revised).ti,ab., 40. patient
safety screener.ti,ab., 41. self inflicted injury severity form.ti,ab., 42. sad persons scale.ti,ab., 43. inter-
sept scale for suicidal thinking.ti,ab., 44. Reasons for living inventory.ti,ab., 45. interpersonal needs
questionnaire.ti,ab., 46. interpersonal needs questionnaire.ti,ab., 47. patient health questionnaire.ti,ab.
48. symptom Driven Diagnostic System for Primary Care.ti,ab., 49. montgomery asberg Depression rat-
ing scale.ti,ab., 50. tool*.ti,ab., 51. instrument*.ti,ab., 54. p4.ti,ab.
52. 1 or 2 or 3 or 4 or 5 or 6, 53. 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17,
55. 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or
35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 54
56. 52 and 53 and 55

COCHRANE LIBRARY

#1(valid*):ti,ab,#2(reliab*):ti,ab, #3(predictive value*):ti,ab, #4(ROC curve):ti,ab., #5(sensitiv*):ti,ab,
#6(specificity*):ti,ab, #7#1 or #2 or #3 or #4 or #5 or #6, #8(suicide*):ti,ab, #9(self killing*):ti,ab, #10(sui-
cidal poisoning):ti,ab, #11(self inflicted death):ti,ab, #12(suicidal behavio?r*):ti,ab, #13(suicid*
risk*):ti,ab, #14(suicid* ideation*):ti,ab, #15(suicide attempt*):ti,ab, #16(attempted suicide*):ti,ab,
#17(parasuicide*):ti,ab, #18(self harm*):ti,ab
#19#8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18
#20(screen* trial*):ti,ab, #21(screen*):ti,ab, #22(questionnaire*):ti,ab, #23(interview*):ti,ab, #24(risk as-
sessment*):ti,ab, #25(risk analysis*):ti,ab, #26(risk evaluation*):ti,ab, #27(safety assessment*):ti,ab,
#28(self report*):ti,ab, #29(diagnos*):ti,ab, #30(diagnos* screen*):ti,ab, #31(tool*):ti,ab, #32(instru-
ment*):ti,ab, #33(suicidal ideation screen*):ti,ab, #34("paykel suicide item"):ti,ab, #35("suicide behaviors
questionnaire"):ti,ab, #36("suicide ideation scale"):ti,ab, #37("suicidal ideation screening question-
naire"):ti,ab, #38("beck depression inventory"):ti,ab, #39("hamilton rating scale for depression"):ti,ab,
#40("beck fast scan"):ti,ab, #41("suicide risk screen"):ti,ab, #42("columbia-suicide severity rating
scale"):ti,ab, #43(("suicide behaviors questionnaire")and revised):ti,ab, #44("patient safety
screener"):ti,ab, #45("self-inflicted injury severity form"):ti,ab, #46("sad persons scale"):ti,ab, #47("inter-
sept scale for suicidal thinking"):ti,ab, #48("reasons for living inventory"):ti,ab, #49("interpersonal needs
questionnaire"):ti,ab, #50("patient health questionnaire"):ti,ab, #51("p4"):ti,ab, #52("symptom driven di-
agnostic system for primary care"):ti,ab, #53("montgomery-asberg depression rating scale"):ti,ab

#54: #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53
#55: #7 and #19 and #54

PSYCHINFO UND PSYINDEX

S9: TI screen* trial* OR AB screen* trial* OR TI screen* OR AB screen* OR TI questionnaire* OR AB questionnaire* OR TI interview* OR AB interview* OR TI risk assessment* OR AB risk assessment* OR TI risk analysis* OR AB risk analysis* OR TI risk evaluation* OR AB risk evaluation* OR TI safety assessment* OR AB safety assessment* OR TI self report* OR AB self report* OR TI diagnos* OR AB diagnos* OR TI diagnos* screen* OR AB diagnos* screen* OR TI tool* OR AB tool* OR TI instrument* OR AB instrument* OR TI suicidal ideation screen* OR AB suicidal ideation screen* OR TI paykel suicide item* OR AB paykel suicide item* OR TI suicide behaviors questionnaire OR AB suicide behaviors questionnaire OR TI suicide ideation scale OR AB suicide ideation scale OR TI suicidal ideation screening questionnaire OR AB suicidal ideation screening questionnaire OR TI beck depression inventory OR AB beck depression inventory OR TI hamilton rating scale for depression OR AB hamilton rating scale for depression OR TI beck fast scan OR AB beck fast scan OR TI suicide risk screen OR AB suicide risk screen OR TI columbia suicide severity rating scale OR AB columbia suicide severity rating scale OR TI (suicide behaviors questionnaire AND revised) OR AB (suicide behaviors questionnaire AND revised) OR TI (suicide behaviours questionnaire AND revised) OR AB (suicide behaviours questionnaire AND revised) OR TI patient safety screener OR AB patient safety screener OR TI suizidalitätsitem aus dem rasch basierten depressionsscreening OR AB suizidalitätsitem aus dem rasch basierten depressionsscreening OR TI self-inflicted injury severity form OR AB self-inflicted injury severity form OR TI sad persons scale OR AB sad persons scale OR TI intersept scale for suicidal thinking OR AB intersept scale for suicidal thinking OR TI reasons for living inventory OR AB reasons for living inventory OR TI interpersonal needs questionnaire OR AB interpersonal needs questionnaire OR TI patient health questionnaire OR AB patient health questionnaire OR TI p4 OR AB p4 OR TI symptom driven diagnostic system for primary care OR AB symptom driven diagnostic system for primary care OR TI montgomery-asberg depression rating scale OR AB montgomeryasberg depression rating scale

S8: TI suicid* risk* OR AB suicid* risk* OR TI suicid* ideation* OR AB suicid* ideation* OR TI suicide attempt* OR AB suicide attempt* OR TI attempted suicide* OR AB attempted suicide* OR TI parasuicide* OR AB parasuicide* OR TI self harm* OR AB self harm* OR TI suicide* OR AB suicide* OR TI self killing* OR AB self killing* OR TI suicidal poisoning OR AB suicidal poisoning OR TI self inflicted death OR AB self inflicted death OR TI suicidal behaviour* OR AB suicidal behaviour* OR TI suicidal behavior* OR AB suicidal behavior*

S7: TI valid* OR AB valid* OR TI reliab* OR AB reliab* OR TI predictive value* OR AB predictive value* OR TI ROC curve OR AB ROC curve OR TI sensitiv* OR AB sensitiv* OR TI specificity* OR AB specificity*

S7 AND S8 AND S9

Extrahierte Items

EXTRAHIERTE ITEMS DER STARD-CHECKLISTE (91)

Scientific and clinical background
Study objectives and hypotheses
Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)
Eligibility criteria
On what basis potentially eligible participants were identified
Where and when potentially eligible participants were identified
Whether participants formed a consecutive, random or convenience series
Index test, in sufficient detail to allow replication
Reference standard, in sufficient detail to allow replication
Rationale for choosing the reference standard (if alternatives exist)
Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory
Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory
Whether clinical information and reference standard results were available to the performers/readers of the index test
Whether clinical information and index test results were available to the assessors of the reference standard
Methods for estimating or comparing measures of diagnostic accuracy
How indeterminate index test or reference standard results were handled
How missing data on the index test and reference standard were handled
Any analyses of variability in diagnostic accuracy
Intended sample size and how it was determined
Flow of participants, using a diagram
Baseline demographic and clinical characteristics of participants
Distribution of severity of disease in those with the target condition
Distribution of alternative diagnoses in those without the target condition
Time interval and any clinical interventions between index test and reference standard
Cross tabulation of the index test results (or their distribution) by the results of the reference standard
Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)
Any adverse events from performing the index test or the reference standard
Study limitations
Implications for practice
Registration number reference
Where the full study protocol can be accessed
Sources of funding and other support

WEITERE EXTRAHIERTE ITEMS

Title
Author
Year
Declaration of interest
Study type
Interview administered or self administered
Instrument freely available

PRISMA-DTA Checklist (79)

Section/topic	PRISMA-DTA Checklist Item	Re-reported on page
TITLE / ABSTRACT		
Title	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	i
Abstract	Abstract: See PRISMA-DTA for abstracts.	v-viii
INTRODUCTION		
Rationale	Describe the rationale for the review in the context of what is already known.	1-7
Clinical role of the index test	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	7-9
Objektives	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	10
METHODS		
Protocol and registration	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	14
Eligibility criteria	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	14, 15
Information sources	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	15, 16
Search	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	16-18, Anhang G
Study selection	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	18, 19
Data collection process	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	20
Definitions for data extraction	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	20
Risk of bias and applicability	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	21, 22
Diagnostic accuracy measures	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	20, 21
Synthesis of results	Describe methods of handling data, combining results of studies and describing variability between studies. This could include but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	23

Meta-analysis	Report the statistical methods used for meta-analyses, if performed.	-
Additional analyses	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	-
RESULTS		
Study selection	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	24, 25
Study characteristics	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	25-34
Risk of bias and applicability	Present evaluation of risk of bias and concerns regarding applicability for each study.	34-37
Results of individual studies	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	37-43
Synthesis of results	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	44, 45
Additional analysis	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	-
DISCUSSION		
Summary of evidence	Summarize the main findings including the strength of evidence.	46-48, 51-58, 60-70 71-73
Limitations	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	48-51, 59, 60 71
Conclusions	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	74, 75
FUNDING		
Funding	For the systematic review, describe the sources of funding and other support and the role of the funders.	-

H LITERATURVERZEICHNIS

1. World Health Organization. Preventing suicide: A global imperative. Geneva: World Health Organization; 2014.
2. O'Connor RC, Nock MK. The psychology of suicidal behaviour. *Lancet Psychiatry*. 2014;1(1):73-85.
3. Mann JJ, Apter A, Bertolote J, Beautrais A, Currier D, Haas A, et al. Suicide prevention strategies: a systematic review. *Jama*. 2005;294(16):2064-74.
4. Forkmann T, Teismann T, Glaesmer H. Diagnostik von Suizidalität. Göttingen: Hogrefe Verlag; 2015.
5. Teismann T, Koban C, Illes F, Oermann A. Psychotherapie suizidaler Patienten. Göttingen: Hogrefe Verlag; 2016.
6. Wolfersdorf M, Etzersdorfer E. Suizid und Suizidprävention. Stuttgart: Kohlhammer Verlag; 2011.
7. Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin. Fachdefinition [cited: 08.08.2019]. Available from: <https://www.degam.de/fachdefinition.html>.
8. Statistisches Bundesamt. Todesursachen, Suizide [cited: 04.12.2021]. Available from: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Todesursachen/Tabellen/suizide.html>.
9. Turecki G, Brent DA. Suicide and suicidal behaviour. *Lancet* 2016;387(10024):1227-39.
10. Hawton K, van Heeringen K. Suicide. *Lancet*. 2009;373(9672):1372-81.
11. Neeleman J, Wessely S. Ethnic minority suicide: a small area geographical study in south London. *Psychol Med*. 1999;29(2):429-36.
12. Bronisch T. Der Suizid: Ursachen, Warnsignale, Prävention. München: CH Beck; 2014.
13. Statistisches Bundesamt. Tabelle Gestorbene: Deutschland, Jahre, Todesursachen, Geschlecht 2016 [cited: 09.07.2019]. Available from: <https://www-genesis.destatis.de/genesis/online/logon?sequenz=tabelleErgebnis&selectionname=23211-0002&sachmerkmal=TODUR1&sachschluessel=TODESURS78&startjahr=1980>.
14. Denning DG, Conwell Y, King D, Cox C. Method choice, intent, and gender in completed suicide. *Suicide & life-threatening behavior*. 2000;30(3):282-8.
15. Erazo N, Baumert JJ, Ladwig K-H. Factors associated with failed and completed railway suicides. *J Affect Disord*. 2005;88(2):137-43.

16. Nock MK, Borges G, Bromet EJ, Alonso J, Angermeyer M, Beautrais A, et al. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br J Psychiatry*. 2008;192(2):98-105.
17. Bernal M, Haro JM, Bernert S, Brugha T, de Graaf R, Bruffaerts R, et al. Risk factors for suicidality in Europe: results from the ESEMED study. *J Affect Disord*. 2007;101(1-3):27-34.
18. Lukaschek K, Engelhardt H, Baumert J, Ladwig K-H. No correlation between rates of suicidal ideation and completed suicides in Europe: Analysis of 49,008 participants (55+ years) based on the Survey of Health, Ageing and Retirement in Europe (SHARE). *Eur Psychiatry*. 2015;30(7):874-9.
19. Schulberg HC, Bruce ML, Lee PW, Williams Jr JW, Dietrich AJ. Preventing suicide in primary care patients: the primary care physician's role. *Gen Hosp Psychiatry*. 2004;26(5):337-45.
20. Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol Bull*. 2017;143(2):187-232.
21. Horwitz AG, Czyz EK, King CA. Predicting future suicide attempts among adolescent and emerging adult psychiatric emergency patients. *J Clin Child Adolesc Psychol*. 2015;44(5):751-61.
22. Brown GK, Beck AT, Steer RA, Grisham JR. Risk factors for suicide in psychiatric outpatients: a 20-year prospective study. *J Consult Clin Psychol*. 2000;68(3):371-7.
23. Bebbington PE, Minot S, Cooper C, Dennis M, Meltzer H, Jenkins R, et al. Suicidal ideation, self-harm and attempted suicide: results from the British psychiatric morbidity survey 2000. *Eur Psychiatry*. 2010;25(7):427-31.
24. Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. *Psychol Med*. 2003;33(3):395-405.
25. Chang EC, Elizabeth AY, Lee JY, Hirsch JK, Kupfermann Y, Kahle ER. An examination of optimism/pessimism and suicide risk in primary care patients: does belief in a changeable future make a difference? *Cognit Ther Res*. 2013;37(4):796-804.
26. Hirsch JK, Conner KR. Dispositional and explanatory style optimism as potential moderators of the relationship between hopelessness and suicidal ideation. *Suicide & life-threatening behavior*. 2006;36(6):661-9.
27. Hirsch JK, Wolford K, LaLonde SM, Brunk L, Morris AP. Dispositional optimism as a moderator of the relationship between negative life events and suicide ideation and attempts. *Cognit Ther Res*. 2007;31(4):533-46.
28. Johnson J, Wood AM, Gooding P, Taylor PJ, Tarrier N. Resilience to suicidality: The buffering hypothesis. *Clin Psychol Rev*. 2011;31(4):563-91.

29. Nock MK, Borges G, Bromet EJ, Cha CB, Kessler RC, Lee S. Suicide and suicidal behavior. *Epidemiol Rev.* 2008;30(1):133-54.
30. O'Connor RC. Towards an integrated motivational–volitional model of suicidal behavior. In: O'Connor R, Platt S, Gordon J, editors. *International Handbook of Suicide Prevention : Research, policy and practice.* Oxford: John Wiley & Sons; 2011. p. 181-98.
31. Baumeister RF. Suicide as escape from self. *Psychol Rev.* 1990;97(1):90-113.
32. Williams M. *Suicide and attempted suicide.* London: Penguin Books; 2001.
33. Joiner T. *Why people die by suicide* Cambridge, Massachusetts, and London, England: Harvard University Press; 2005.
34. Rudd MD. Fluid Vulnerability Theory: A Cognitive Approach to Understanding the Process of Acute and Chronic Suicide Risk. In: Ellis TE, editor. *Cognition and suicide: Theory, research, and therapy.* Washington: American Psychological Association; 2006. p. 355-68.
35. Wenzel A, Beck AT. A cognitive model of suicidal behavior: Theory and treatment. *Appl Prev Psychol.* 2008;12(4):189-201.
36. Zalsman G, Hawton K, Wasserman D, van Heeringen K, Arensman E, Sarchiapone M, et al. Suicide prevention strategies revisited: 10-year systematic review. *Lancet Psychiatry.* 2016;3(7):646-59.
37. DGPPN, BÄK, KBV, AWMF. S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression – Kurzfassung 2017 [cited: 28.02.2019]. Available from: www.depression.versorgungsleitlinien.de.
38. Excellence NIfC. Quick reference guide - Depression: management of depression in primary and secondary care: National Institute for Clinical Excellence; 2004 [cited: 07.08.2019]. Available from: <http://www.scamfyc.org/documentos/depression%20NICE.pdf>.
39. Gould MS, Greenberg T, Velting DM, Shaffer D. Youth suicide risk and preventive interventions: a review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2003;42(4):386-405.
40. Meehan J, Kapur N, Hunt IM, Turnbull P, Robinson J, Bickley H, et al. Suicide in mental health in-patients and within 3 months of discharge: National clinical survey. *British Journal of Psychiatry.* 2018;188(2):129-34.
41. Pirkis J, Blood RW, Beautrais A, Burgess P, Skehans J. Media guidelines on the reporting of suicide. *Crisis.* 2006;27(2):82-7.
42. Nationales Suizid Präventions Programm. Das Nationale Suizidpräventionsprogramm für Deutschland (NaSPro) [cited: 04.12.2021]. Available from: <https://www.suizidpraevention-deutschland.de>.

43. Deutsche Gesellschaft für Suizidprävention. Startseite [cited: 5.07.2019]. Available from: <https://www.suizidprophylaxe.de>.
44. Schneider B, Lindner R, Giegling I, Müller S, Müller-Pein H, Rujescu D, et al. Suizidprävention Deutschland. Kassel 2021 [cited: 04.12.2021]. Available from: <https://www.naspro.de/dl/Suizidpraevention-Deutschland-2021.pdf>.
45. Luoma JB, Martin CE, Pearson JL. Contact with mental health and primary care providers before suicide: a review of the evidence. *American Journal of Psychiatry*. 2002;159(6):909-16.
46. Stene-Larsen K, Reneflot A. Contact with primary and mental health care prior to suicide: a systematic review of the literature from 2000 to 2017. *Scand J Public Health*. 2019;47(1):9-17.
47. Hauge LJ, Stene-Larsen K, Grimholt TK, Øien-Ødegaard C, Reneflot A. Use of primary health care services prior to suicide in the Norwegian population 2006-2015. *BMC Health Serv Res*. 2018;18(1):619-.
48. Pearson A, Saini P, Da Cruz D, Miles C, While D, Swinson N, et al. Primary care contact prior to suicide in individuals with mental illness. *Br J Gen Pract*. 2009;59(568):825-32.
49. Houston K, Haw C, Townsend E, Hawton K. General practitioner contacts with patients before and after deliberate self harm. *Br J Gen Pract*. 2003;53(490):365-70.
50. Isometsa ET, Heikkinen ME, Marttunen MJ, Henriksson MM, Aro HM, Lonnqvist JK. The last appointment before suicide: is suicide intent communicated? *The American journal of psychiatry*. 1995;152(6):919-22.
51. Lukaschek K, Erazo N, Baumert J, Ladwig K-H. Suicide mortality in comparison to traffic accidents and homicides as causes of unnatural death. An analysis of 14,441 cases in Germany in the year 2010. *Int J Environ Res Public Health*. 2012;9(3):924-31.
52. Gensichen J, Teising A, König J, Gerlach FM, Petersen JJ. Predictors of suicidal ideation in depressive primary care patients. *J Affect Disord*. 2010;125(1-3):124-7.
53. Teismann T, Lukaschek K, Hiller TS, Breitbart J, Brettschneider C, Schumacher U, et al. Suicidal ideation in primary care patients suffering from panic disorder with or without agoraphobia. *BMC Psychiatry*. 2018;18(1):305.
54. Kruse W, Saß H. Curriculum Psychosomatische Grundversorgung-Basisdiagnostik und Basisversorgung bei Patienten mit psychischen und psychosomatischen Störungen einschließlich Aspekte der Qualitätssicherung. Köln: Bundesärztekammer 2001 [cited: 06.08.2019]. Available from: https://www.aerzte-seminar.de/wpcontent/uploads/2018/07/Curriculum_PsychosomatischeGrundversorgung_BAEK.pdf.

55. DEGAM. Kompetenzbasiertes Curriculum Allgemeinmedizin 2015 [cited: 19.07.2019]. Available from: https://www.weiterbildung-allgemeinmedizin.de/downloads/Curriculum_01-10-15.pdf.
56. Kliem S, Lohmann A, Mößle T, Brähler E. German Beck scale for suicide ideation (BSS): psychometric properties from a representative population survey. *BMC Psychiatry*. 2017;17(1):389.
57. Ressing M, Blettner M, Klug SJ. Systematische Übersichtsarbeiten und Metaanalysen. *Dtsch Arztebl Int*. 2009;106(27):456-63.
58. Murad MH, Montori VM, Ioannidis JP, Jaeschke R, Devereaux P, Prasad K, et al. How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *Jama*. 2014;312(2):171-9.
59. Mallett S, Halligan S, Thompson M, Collins GS, Altman DG. Interpreting diagnostic accuracy studies for patient care. *BMJ*. 2012;345:e3999.
60. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working G. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149(12):889-97.
61. Glasziou P, Irwig L, Bain C, Colditz G. Systematic reviews in health care: a practical guide. Cambridge: Cambridge University Press; 2001.
62. Cochrane Collaboration. Cochrane handbook for systematic reviews of diagnostic test accuracy [cited: 04.12.2021]. Available from: <https://training.cochrane.org/resource/cochrane-handbook-systematic-reviews-diagnostic-test-accuracy>.
63. Brown GK. A review of suicide assessment measures for intervention research with adults and older adults. 2001.
64. Runeson B, Odeberg J, Pettersson A, Edbom T, Adamsson IJ, Waern M. Instruments for the assessment of suicide risk: a systematic review evaluating the certainty of the evidence. *PLoS ONE*. 2017;12(7):e0180292.
65. Batterham PJ, Ftanou M, Pirkis J, Brewer JL, Mackinnon AJ, Beautrais A, et al. A systematic review and evaluation of measures for suicidal ideation and behaviors in population-based research. *Psychol Assess*. 2015;27(2):501-12.
66. Abarca C, Gheza C, Coda C, Elicer B. Literature review to identify standardized scales for assessing adult suicide risk in the primary health care setting. *Medwave*. 2018;18(5):e7246.
67. Kreuze E, Lamis DA. A review of psychometrically tested instruments assessing suicide risk in adults. *Omega (Westport)*. 2018;77(1):36-90.
68. Roos L, Sareen J, Bolton JM. Suicide risk assessment tools, predictive validity findings and utility today: time for a revamp? *Neuropsychiatry (London)*. 2013;3(5):1-13.

69. Cassidy S, Bradley L, Bowen E, Wigham S, Rodgers J. Measurement properties of tools used to assess suicidality in autistic and general population adults: a systematic review. *Clin Psychol Rev.* 2018;62:56-70.
70. Range LM, Knott EC. Twenty suicide assessment instruments: Evaluation and recommendations. *Death Stud.* 1997;21(1):25-58.
71. Horowitz LM, Ballard ED, Pao M. Suicide screening in schools, primary care and emergency departments. *Curr Opin Pediatr.* 2009;21(5):620-7.
72. Range LM. The family of instruments that assess suicide risk. *J Psychopathol Behav Assess.* 2005;27(2):133-40.
73. Gaynes BN, West SL, Ford CA, Frame P, Klein J, Lohr KN. Screening for suicide risk in adults: a summary of the evidence for the US Preventive Services Task Force. *Ann Intern Med.* 2004;140(10):822-35.
74. O'Connor E, Gaynes BN, Burda BU, Soh C, Whitlock EP. Screening for and treatment of suicide risk relevant to primary care: a systematic review for the US Preventive Services Task Force. *Ann Intern Med.* 2013;158(10):741-54.
75. Brähler E, Zenger M, Kemper CJ. Psychologische und sozialwissenschaftliche Kurzskalen: Standardisierte Erhebungsinstrumente für Wissenschaft und Praxis. Berlin: MWV; 2015.
76. National Institute for Health Research, University of York Centre for Reviews and Dissemination. PROSPERO International prospective register of systematic reviews [cited: 04.06.2019]. Available from: <https://www.crd.york.ac.uk/prospero/>.
77. Altman DG, Simera I, Hoey J, Moher D, Schulz K. EQUATOR: reporting guidelines for health research. *Lancet.* 2008;371(9619):1149-50.
78. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med.* 2009;6(7):e1000100.
79. McInnes MD, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *Jama.* 2018;319(4):388-96.
80. Müllner M. Erfolgreich wissenschaftlich arbeiten in der Klinik: evidence based medicine. Wien: Springer-Verlag; 2005.
81. Frank M, Lukaschek K, Leucht S, Gensichen J. A systematic review of brief screeners for suicidal behaviour in primary care. PROSPERO. 2019.
82. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club.* 1995;123(3):A12-3.

83. Lukaschek K, Frank M, Halfter K, Schneider A, Gensichen J. Kurze Screener für Suizidalität in der Allgemeinmedizin. *MMW - Fortschritte der Medizin*. 2021;163(6):9-18.
84. Wolters Kluwer. Ovid 2019 [cited: 06.08.2019]. Available from: <http://ovidsp.dc2.ovid.com.emedien.ub.uni-muenchen.de>.
85. Cochrane Collaboration. Cochrane Library 2019 [cited: 13.06.2019]. Available from: <https://www-cochranelibrary-com.emedien.ub.uni-muenchen.de/>.
86. EBSCO Information Services GmbH. EBSCO 2019 [cited: 13.06.2019]. Available from: <http://web.a.ebscohost.com.emedien.ub.uni-muenchen.de/>.
87. King CA, Horwitz A, Czyz E, Lindsay R. Suicide risk screening in healthcare settings: Identifying males and females at risk. *J Clin Psychol Med Settings*. 2017;24(1):8-20.
88. Dube P, Kurt K, Bair MJ, Theobald D, Williams LS. The p4 screener: evaluation of a brief measure for assessing potential suicide risk in 2 randomized effectiveness trials of primary care and oncology patients. *Prim Care Companion J Clin Psychiatry*. 2010;12(6):e1-e8.
89. O'Rourke M, Jamil RT, Siddiqui W. *Suicide Screening and Prevention*. Treasure Island (FL): StatPearls Publishing LLC.; 2019 [cited: 06.08.2019]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK531453/>.
90. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev*. 2016;5(1):210.
91. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology*. 2015;277(3):826-32.
92. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem*. 2003;49(1):7-18.
93. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-36.
94. University of Bristol. QUADAS-2 [cited: 12.06.2019]. Available from: <https://www.bristol.ac.uk/population-health-sciences/projects/quadas/quadas-2/>.
95. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-94.
96. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008;336(7653):1106-10.

97. McMaster University. GRADEpro Guideline Development Tool [Software] 2015 [cited: 06.08.2019]. Available from: grade.pro.org.
98. Schünemann HJ. Introductory courses for GRADE and Summary of Findings Tables - Inconsistency 2011 [cited: 13.06.2019]. Available from: <https://cebgrade.mcmaster.ca/Inconsistency/index.html>.
99. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: Analysing and Presenting Results. 2010 [cited 09.08.2019]. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 10 [Internet]. The Cochrane Collaboration, [cited 09.08.2019]. Available from: <http://srdta.cochrane.org/>.
100. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264-9.
101. Cooper-Patrick L, Crum RM, Ford DE. Identifying suicidal ideation in general medical patients. *Jama*. 1994;272(22):1757-62.
102. Olfson M, Weissman MM, Leon AC, Sheehan DV, Farber L. Suicidal ideation in primary care. *J Gen Intern Med*. 1996;11(8):447-53.
103. Uebelacker LA, German NM, Gaudiano BA, Miller IW. Patient health questionnaire depression scale as a suicide screening instrument in depressed primary care patients: a cross-sectional study. *Prim Care Companion CNS Disord*. 2011;13(1):e1-e6.
104. Chamberlain P, Goldney R, Delfabbro P, Gill T, Dal Grande L. Suicidal ideation: the clinical utility of the K10. *Crisis*. 2009;30(1):39-42.
105. Millner AJ, Lee MD, Nock MK. Single-item measurement of suicidal behaviors: Validity and consequences of misclassification. *PLoS ONE*. 2015;10(10):e0141606.
106. Van Spijker BA, Batterham PJ, Caelear AL, Farrer L, Christensen H, Reynolds J, et al. The Suicidal Ideation Attributes Scale (SIDAS): Community-based validation study of a new scale for the measurement of suicidal ideation. *Suicide & life-threatening behavior*. 2014;44(4):408-19.
107. Caelear AL, Batterham PJ, Sunderland M, Carragher N. Development and validation of static and adaptive screeners to assess suicidal thoughts and behavior. *Suicide and Life-Threatening Behavior*. 2020;50(1):189-200.
108. Taylor A, Dal Grande E, Gill T, Fisher L, Goldney R. Detecting determinants of suicidal ideation: South Australian surveillance system results. *Int J Public Health*. 2007;52(3):142-52.
109. Kessler R, Mroczek D. Final versions of our non-specific psychological distress scale. University of Michigan. Ann Arbor: Survey Research Center of the Institute for Social Research. 1994.
110. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13.

111. Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. *Psychol Med.* 1979;9(1):139-45.
112. Watson D, Goldney R, Fisher L, Merritt M. The measurement of suicidal ideation. *Crisis.* 2001;22(1):12-4.
113. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia–Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *The American journal of psychiatry.* 2011;168(12):1266-77.
114. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule: Its History, Characteristics, and Validity. *JAMA Psychiatry.* 1981;38(4):381-9.
115. First MB, Spitzer RL, Gibbon M, Williams JB. Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. SCID-I/P New York, NY; 2002.
116. University of Bristol. Tabular Display [cited: 18.06.2019]. Available from: <https://www.bristol.ac.uk/population-health-sciences/projects/quadas/resources/>.
117. University of Bristol. Graphical display [cited: 18.06.2019]. Available from: <https://www.bristol.ac.uk/population-health-sciences/projects/quadas/resources/>.
118. Gerhardt C. Ueber Diagnostik und Therapie. *Dtsch Med Wochenschr.* 1885;11(45):778-80.
119. Joint Commission. Detecting and treating suicide ideation in all settings. *Sentinel Event Alert.* 2016(56):1-7.
120. Spix C, Blettner M. Screening: part 19 of a series on evaluation of scientific publications. *Dtsch Arztebl Int.* 2012;109(21):385.
121. Hom MA, Joiner Jr TE, Bernert RA. Limitations of a single-item assessment of suicide attempt history: Implications for standardized suicide risk assessment. *Psychol Assess.* 2016;28(8):1026.
122. Gardner DG, Cummings LL, Dunham RB, Pierce JL. Single-item versus multiple-item measurement scales: An empirical comparison. *Educ Psychol Meas.* 1998;58(6):898-915.
123. Simon GE, Rutter CM, Peterson D, Oliver M, Whiteside U, Operskalski B, et al. Does response on the PHQ-9 Depression Questionnaire predict subsequent suicide attempt or suicide death? *Psychiatric services (Washington, DC).* 2013;64(12):1195-202.
124. The Assessment and Management of Suicide Risk Work Group. VA/DoD Clinical Practice Guideline for the Assessment and Management of Patients at Risk for Suicide: Department of Veterans Affairs

Department of Defense; 2019 [cited: 28.08.2019]. Available from: <https://www.healthquality.va.gov/guidelines/MH/srb/VADoDSuicideRiskFullCPGFinal5088212019.pdf>.

125. Thompson AH, Dewa CS, Phare S. The suicidal process: age of onset and severity of suicidal behaviour. *Social psychiatry and psychiatric epidemiology*. 2012;47(8):1263-9.

126. Michigan Quality Improvement Consortium Guideline. Michigan Quality Improvement Consortium Guideline Primary Care Diagnosis and Management of Adults with Depression 2018 [cited: 16.07.2019]. Available from: http://mqic.org/pdf/mqic_primary_care_diagnosis_and_management_of_adults_with_depression_cpg.pdf.

127. Gould MS, Marrocco FA, Kleinman M, Thomas JG, Mostkoff K, Cote J, et al. Evaluating iatrogenic risk of youth suicide screening programs: a randomized controlled trial. *Jama*. 2005;293(13):1635-43.

128. Harris KM, Goh MT. Is suicide assessment harmful to participants? Findings from a randomized controlled trial. *Int J Ment Health Nurs*. 2017;26(2):181-90.

129. Teismann T, Dorrman W. Suizidalität. Göttingen: Hogrefe Verlag; 2014.

130. Gallagher MW, Lopez SJ. Positive expectancies and mental health: Identifying the unique contributions of hope and optimism. *J Posit Psychol*. 2009;4(6):548-56.

131. Jobes DA, Jacoby AM, Cimboic P, Hustead LAT. Assessment and treatment of suicidal clients in a university counseling center. *J Couns Psychol*. 1997;44(4):368.

132. Joiner Jr TE, Rudd MD, Rajab MH. Agreement between self-and clinician-rated suicidal symptoms in a clinical sample of young adults: Explaining discrepancies. *J Consult Clin Psychol*. 1999;67(2):171-6.

133. Trivedi MH, Wisniewski SR, Morris DW, Fava M, Gollan JK, Warden D, et al. Concise Health Risk Tracking scale: a brief self-report and clinician rating of suicidal risk. *J Clin Psychiatry*. 2011;72(6):757-64.

134. McNamee JE, Offord DR. Prevention of Suicide In: The Canadian Task Force on the Periodic Health Examination, editor. *The Canadian guide to clinical preventive health care* Ottawa: Canada Communication Group-Publishing; 1994. p. 456-67.

135. Attia J. Moving beyond sensitivity and specificity: using likelihood ratios to help interpret diagnostic tests. *Aust Prescr*. 2003;26(5):111-13.

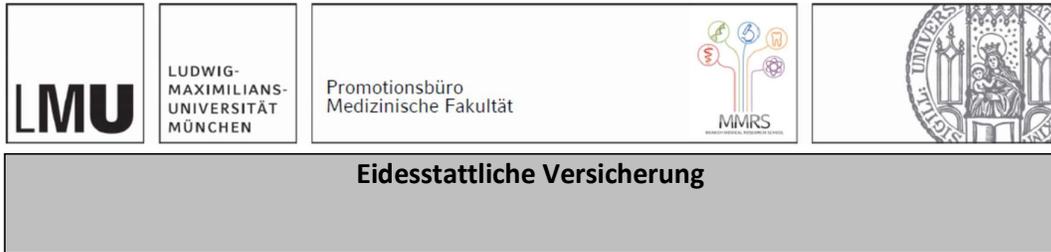
136. Rosen A. Detection of suicidal patients: an example of some limitations in the prediction of infrequent events. *J Consult Psychol*. 1954;18(6):397-403.

137. Pokorny AD. Prediction of suicide in psychiatric patients: report of a prospective study. *Arch Gen Psychiatry*. 1983;40(3):249-57.

138. Carter G, Milner A, McGill K, Pirkis J, Kapur N, Spittal MJ. Predicting suicidal behaviours using clinical instruments: systematic review and meta-analysis of positive predictive values for risk scales. *Br J Psychiatry*. 2017;210(6):387-95.
139. Osman A, Bagge CL, Gutierrez PM, Konick LC, Kopper BA, Barrios FX. The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment*. 2001;8(4):443-54.
140. Giddens JM, Sheehan KH, Sheehan DV. The Columbia-Suicide Severity Rating Scale (C-SSRS): Has the "Gold Standard" Become a Liability? *Innov Clin Neurosci*. 2014;11(9-10):66-80.
141. Harris KM, Syu J-J, Lello OD, Chew YE, Willcox CH, Ho RH. The ABC's of suicide risk assessment: Applying a tripartite approach to individual evaluations. *PLoS One*. 2015;10(6):e0127442.
142. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). Guidance for Industry: Suicidal Ideation and Behavior: Prospective Assessment of Occurrence in Clinical Trials 2012 [cited: 07.08.2019]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-suicidal-ideation-and-behavior-prospective-assessment-occurrence-clinical-trials#iii_pROSPECTIVE.
143. Batterham PJ, Calear AL, van Spijker BA. The specificity of the interpersonal-psychological theory of suicidal behavior for identifying suicidal ideation in an online sample. *Suicide & life-threatening behavior*. 2015;45(4):448-60.
144. Turnell AI, Fassnacht DB, Batterham PJ, Calear AL, Kyrios M. The Self-Hate Scale: Development and validation of a brief measure and its relationship to suicidal ideation. *J Affect Disord*. 2019;245:779-87.
145. Fountoulakis KN, Pantoula E, Siamouli M, Moutou K, Gonda X, Rihmer Z, et al. Development of the Risk Assessment Suicidality Scale (RASS): a population-based study. *J Affect Disord*. 2012;138(3):449-57.
146. Rimkeviciene J, Hawgood J, O'Gorman J, De Leo D. Construct validity of the acquired capability for suicide scale: factor structure, convergent and discriminant validity. *J Psychopathol Behav Assess*. 2017;39(2):291-302.
147. Glaesmer H, Kapusta ND, Teismann T, Wagner B, Hallensleben N, Spangenberg L, et al. Psychometrische Eigenschaften der deutschen Version des Suicide Behaviors Questionnaire Revised (SBQ-R). *Psychother Psychosom Med Psychol*. 2018;68(08):346-52.
148. Bernau R, Börger M, Egidi G, Mühlendorf H-M, Schmiemann G. Kriterienkatalog für die Weiterbildungsbefugnis für die Facharztweiterbildung Allgemeinmedizin 2016 [cited: 23.08.2019]. Available from: [https://www.degam.de/files/Inhalte/Degam-Inhalte/Sektionen und Arbeitsgruppen/Sektion Fortbildung/Entscheidungshilfe Weite rbildungsbefugnis Allgemeinmedizin Ueberarbeitung 2016 4.pdf](https://www.degam.de/files/Inhalte/Degam-Inhalte/Sektionen%20und%20Arbeitsgruppen/Sektion%20Fortbildung/Entscheidungshilfe%20Weiterbildungsbefugnis%20Allgemeinmedizin%20Ueberarbeitung%202016%204.pdf).

149. Richards J, Parrish R, Lee A, Bradley K, Caldeiro R. An Integrated Care Approach to Identifying and Treating the Suicidal Person in Primary Care Psychiatr Times. 2019;36(1):1-4.
150. LeFevre ML. Screening for suicide risk in adolescents, adults, and older adults in primary care: US Preventive Services Task Force recommendation statement. Ann Intern Med. 2014;160(10):719-26.
151. Weckbecker K, Schläpfer TE. Hausärzte können Suizid verhindern 2013 [cited: 23.08.2019]. Available from: <https://www.uni-bonn.de/neues/056-2013>.
152. Suicide Assessment Five-step Evaluation and Triage for Mental Health Professionals. 2009 [cited: 29.08.2019]. Available from: https://www.integration.samhsa.gov/images/res/SAFE_T.pdf.
153. Bauer AM, Chan Y-F, Huang H, Vannoy S, Unützer J. Characteristics, management, and depression outcomes of primary care patients who endorse thoughts of death or suicide on the PHQ-9. Journal of general internal medicine. 2013;28(3):363-9.
154. Ahmedani BK, Simon GE, Stewart C, Beck A, Waitzfelder BE, Rossom R, et al. Health care contacts in the year before suicide death. Journal of general internal medicine. 2014;29(6):870-7.

I AFFIDAVIT



Eidesstattliche Versicherung

Frank, Milena

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Kurze Instrumente zum Screenen von Suizidalität in der Hausarztpraxis:
Ein systematischer Review

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den 15.01.2023

Milena Frank

J CURRICULUM VITAE

Milena Frank

Geboren am 13.03.1992
in München

Berufserfahrung

Seit März 2020 Weiterbildung zur Fachärztin für Augenheilkunde

Studium

*Oktober 2018 –
Dezember 2022* Promotion am Institut für Allgemeinmedizin der LMU

Juli 2018 Erteilung der Approbation als Ärztin

*Oktober 2013 -
Mai 2018* Studium der Humanmedizin an der TUM
Zweiter Abschnitt der Ärztlichen Prüfung und Dritter Abschnitt
der Ärztlichen Prüfung

*Oktober 2011 -
September 2013* Studium der Humanmedizin an der LMU
Erster Abschnitt der Ärztlichen Prüfung

*Oktober 2014 -
September 2015* Parallelstudium der Zahnmedizin an der LMU
Abschluss der Vorklinik

Schulausbildung

*September 2002 -
Mai 2011* Besuch des Käthe-Kollwitz-Gymnasiums München
Abitur