A systematic review of brief screeners for suicidal behaviour in primary care

CURRENT STATUS: POSTED

Karoline Lukaschek
Institute of General Practice and Family Medicine, University Hospital of the Ludwig-Maximilians Universitaet Muenchen

Milena Frank
Institute of General Practice and Family Medicine, University hospital of the Ludwig-Maximilians-Universitaet Muenchen

Kathrin Halfter
Munich Cancer Registry at the University Hospital of Munich, Insitute for medical Information Processing, Biometry and Epidemiology, Ludwig-Maximilians-University

Antonius Schneider
Institute of General Practice, University Hospital Klinikum rechts der Isar, Technische Universität Muenchen

Jochen Gensichen
Institute of General Practice and Family Medicine, University Hospital of the Ludwig-Maximilians-Universitaet Muenchen

Email: allgemeinmedizin@med.uni-muenchen.de
ORCiD: 0000-0003-0644-8597

DOI: 10.21203/rs.2.14068/v1

SUBJECT AREAS
Psychiatry

KEYWORDS
Suicide prevention, screening, diagnostic accuracy, primary care, general population
Abstract

Background As primary contacts, general practitioners can play a pivotal role in identifying suicidal behaviour in their patients. Therefore, actively screening patients at risk should be an integral part of primary care services. We summarized the available studies reporting diagnostic accuracy of short screening instruments for suicidal behaviour in primary care or the general population in a narrative synthesis. A brief instrument could improve the identification of vulnerable patients and their subsequent referral to mental health specialist.

Methods: The following databases MEDLINE, EMBASE, PsychINFO, PSYNDEX, and Cochrane Library were searched in January 2019 without any time constraints. Risk of bias and applicability concerns were assessed using the QUADAS-2 tool. The certainty of evidence was rated via GRADEpro. The authors followed the PRISMA extensions for Diagnostic Test Accuracy Studies.

Results Six relevant studies fulfilled all criteria and were included. Each index test was only assessed in a single study. The studies each used different reference standards. Target conditions defined across the different studies were suicidal ideation, suicide plan, suicide attempt, and suicidal behaviour in general. The diagnostic accuracy measurements sensitivity and specificity had a wide range (sensitivity: 26% - 100%, specificity: 64% - 99%). Risk of bias was rated moderate and concerns regarding applicability acceptable. A required sensitivity of at least 80% and specificity of 50% with a moderate to high GRADE rating was achieved by six of nine index tests.

Conclusions Overall, only a few studies assessed screening instruments for suicidal behaviour in primary care or the general population. The identified studies were heterogeneous regarding sample size, index test, and reference standard. Since each index test was only evaluated in one study, their interpretability is limited. No uniform
gold standard was found, as the studies identified used different reference standards. Although there are other promising short questionnaires, the majority has not yet been evaluated in primary care setting or the general population with regard to their diagnostic accuracy. A final assessment should always be based on the clinical judgement of the attending physician. Registration The study protocol was registered at PROSPERO (ID: CRD42019122173).

Background

Suicidal behaviour is described as a continuous process. It begins with occasional suicidal thoughts and can lead to the complete suicide (1). A hierarchical arrangement of different degrees of severity also forms the basis of the progression from suicidal thoughts to suicide attempts (2). We defined suicidal behaviour in accordance with O’Conner et al. (2014): suicidal behaviour refers to “thoughts and behaviours related to an individual intentionally taking their own life. These thoughts include the more specific outcomes of suicide ideation, which refers to an individual having thoughts about intentionally taking their own life; suicide plan, which refers to the formulation of a specific plot by an individual to end their own life; and suicide attempt, which refers to engagement in a potentially self-injurious behaviour in which there is at least some intention of dying as a result of the behaviour. “(3).

In the general population, lifetime prevalence of suicidal ideation ranges between 7.8 and 9% (4, 5), short-time or point-prevalence are usually lower (6–8). The prevalence of suicide plan and suicide attempt was approximately 3% (9). About 30% of the individuals with suicidal ideation attempt suicide - two thirds of which in the first year after the onset of suicidal ideation (9). Studies of suicidal behaviour in primary care showed that around 1% to 10% of the patients had suicidal thoughts (10), but only a few of them talked about these thoughts to their general practitioner (GP) (11–13). Suicidal ideation and previous
suicidal attempts are considered strong predictors of future suicide (14–17). Therefore, actively screening for suicidal behaviour should be an integral part of primary care services: it could reach large numbers of individuals, especially at high-risk patients, who otherwise would not be identified (18).

The literature shows, that up to 80% of suicide victims visited their GP in the year prior to their suicide and almost half even a month before (19, 20). However, GPs often did not recognize the symptoms of a potential suicide risk in their patients’ behaviour (21, 22). Suicide prevention strategies in primary care aim at training and increasing the awareness of GPs. Programs include information on the treatment of mental disorders (23) and recognizing patients with suicidal behaviour through screening (18). In 2013, the U.S. Preventive Services Task Force published a review on the subject of screening for suicide risk in primary care (24). The review showed limited data indicating that there are screening instruments with acceptable performance characteristics for adults (24). The present review also focuses on screening instruments for suicidal patients in primary care. In comparison to the Task force review, several additional criteria were taken into account to broaden the scope of insight on this subject: studies published after 2013 were added to our search strategy; to our knowledge, this is the first review on this subject also including articles in German language; furthermore, the number of items of the included screening instruments was limited to a maximum of twelve items (25). Due to the small number of primary care studies to be expected from preliminary literature search results, we also included studies from the general population with a similarly low prevalence of suicidal behaviour (9) (10). The aim of this study was to obtain an initial overview of the available screening instruments and compare their capability in correctly identifying patients with suicidal behaviour and in need of further mental health care.

Methods
The authors followed the PRISMA extensions for Diagnostic Test Accuracy Studies (26).

Protocol and registration

The study protocol was registered at the PROSPERO international prospective register of systematic reviews (ID: CRD42019122173).

Eligibility criteria

We defined the following eligibility criteria for studies to be included in the present review. Participants had to be primary care patients or participants from the general population, aged 14 years or older, and predominantly Caucasian. Suicidal behaviour was the defined target condition and splits itself according to the definition of suicidal behaviour from O’ Conner et al. (2014) (3) into different aspects. These are suicidal ideation, suicide plan, suicide attempt, and suicidal behaviour in general (see background). Studies were included if the index test had up to twelve items and was used as a screener for suicidal behaviour. The reference standard should be able to measure suicidal behaviour as a comparator. Both tests could be administered in the form of an interview or as a self-administered questionnaire. The different screening instruments were compared with regard to their diagnostic accuracy in correctly identifying patients at risk for suicidal behaviour. We included diagnostic studies and randomised controlled trials published in English or German language.

Information sources

A skilled librarian provided assistance for the development of the electronic search. A keyword search in title and abstract and a thesaurus search was performed in the following five databases: MEDLINE (Ovid) (27), EMBASE (Ovid) (27), Cochrane Library (Wiley) (28), PSYNDEx (EBSCO) (29), and PsychINFO (EBSCO) (29). The search was performed in January 2019; no time limits were used. The full search strategy is shown in
the supplementary file 1. Through an additional manual search further articles were identified. This included screening references of included studies, identified reviews, meta-analyses, and books.

**Study selection**

For deduplication all search results were exported to EndNote. Afterwards, two reviewers Karoline Lukaschek (KL) and Milena Frank (MF) screened titles and abstracts independently using the web application Rayyan (30). Studies for the full-text analysis were selected and checked for eligibility by two reviewers (KL, MF). Studies fulfilling eligibility criteria were included in the review. Any disagreements were resolved by consensus. If this did not lead to an agreement, a third reviewer Jochen Gensichen (JG) was consulted.

**Data collection process**

Both reviewers (KL, MF) performed data collection for the key data points. All relevant data was extracted and added to a Microsoft Access sheet, which was prepared for this review. Any disagreements between the two reviewers in the data collection were resolved by discussion among the researchers. If this did not lead to an agreement, a third reviewer (JG) was again consulted.

For data extraction a part of the STARD 2015 checklist was used (31). Key characteristics of the studies such as participant characteristics, clinical setting, study design, target condition, index test, reference standard, sample size, and funding were extracted. Corresponding authors were contacted if necessary.

**Risk of bias and applicability**

Risk of bias and applicability was assessed using Quality Assessment of Diagnostic
Accuracy Studies—version 2 (QUADAS–2) by KL and MF (32). In order to evaluate the risk of bias for each study, different categories were assessed. These included patient selection, index test, reference standard and flow and timing. To estimate the applicability concerns, patient selection, index test and reference standard were judged. The different categories were classified as low, high or unclear. The quality assessment was conducted by two independent reviewers and any disagreements were resolved by consensus or by consultation of JG.

Diagnostic accuracy measures

The different screeners were assessed and compared with regard to their measures of diagnostic accuracy. This included sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). These measures were extracted from original studies if reported or calculated with ThresholdROC package version 2.7 in R version 3.6.0 by Kathrin Halfter (KH) and MF. If not available, 95% confidence intervals were calculated using Wilson Score interval (KH and MF). Sensitivity and specificity were visualized in a forest plot (see Figure 2).

Certainty of evidence

The certainty of evidence of the different index tests was assessed using GRADEpro. The certainty of evidence was classified as high, moderate, low or very low. Different factors that may decrease the certainty of evidence were evaluated. This included Risk of bias, Indirectness, Inconsistency, Imprecision, and Other considerations (33). If there was only a single study for a screening instrument, we followed the recommendation of the Cochrane Training and did not downgrade in relation to Inconsistency (34).

Synthesis of results

Due to the large heterogeneity of the extracted studies a narrative synthesis combined
with figures and tables was performed. Sensitivity and specificity were visualized in a forest plot.

Results

Study selection

**Figure 1 PRISMA Flow Diagram (35)**

A total of 9,969 results were identified through database searching and additional search. After removing duplicates, title and abstract of the remaining 5,324 studies were screened. Of these, 55 articles were included in the full text analysis. A total of 49 articles were excluded with reasons (see Figure 1). Short screening instruments for suicidal behaviour in primary care or the general population were evaluated in 14 articles (36-49). In these studies, over 17 different screening instruments for suicidal behaviour were assessed. Including Interpersonal-Psychological Theory of Suicidal Behaviour (42), P4 (48), Risk Assessment Suicidality Scale (RASS) (45), Suicide Behaviours Questionnaire Revised (SBQ-R) (49), Suicidal Affect-Behaviour-Cognition Scale (SABCS) (47), German Beck Scale for Suicide Ideation (BSS) (43), Self-Hate Scale (44), Acquired Capability for Suicide Scale (ACSS) (46), Kessler Psychological Distress Scale (K10) (39), Suicidal Ideation Screening Questionnaire (SIS-Q) (40), Suicidal Ideation Attributes Scale (SIDAS) (36), as well as the single item measures Gate question suicide attempt (37), Gate question suicide ideation (37), Feeling suicidal, Wishing you were dead (38), Thoughts of death, and Patient-Health-Questionaire–9 (PHQ–9) - item 9 (41).

Only n = 6 of the identified articles measured diagnostic test accuracy or provided data for calculating it (36-41) (see Table 1). These were included in the narrative synthesis. Table 2 gives an overview of the remaining n = 9 studies not further assessed, due to missing diagnostic accuracy measurements (42-49).
Study characteristics

The six included studies all measured different index tests and used various reference standards. Four studies were conducted in the United States (37, 38, 40, 41) and two in Australia (36, 39). There was no study identified in German language. The oldest study dates back to 1994 (40) and the most recent one was published in 2015 (37). All studies had a cross-sectional design. Three articles evaluated suicidal behaviour in primary care (38, 40, 41) and three in the general population (36, 37, 39). Suicidal ideation was the most widely used target condition (four times), followed by suicide attempt (two times) and suicide plan (two times). The general term suicidal behaviour was selected only once as target condition. Each of the described index tests were assessed only in a single study. Table 1 shows key characteristics for included studies.

Risk of bias and applicability

Regarding patient selection, 33% of the articles were assessed as having low risk of bias, 33% were judged as high risk of bias, and 33% as unclear risk of bias. The index test was assessed 83% as low risk of bias and 17% as unclear risk of bias. The category reference standard was assessed 33% as low risk of bias and 67% as unclear risk of bias. Flow and timing were judged 50% as low risk of bias, 33% as unclear of bias and 17% as high risk of bias. The applicability concerns were rated as low risk, just concerning the patient selection, applicability concerns were unclear risk in 33% of the articles.

For further information, see Table 3 Quadas–2 tool (32).

Results of individual studies and synthesis of results

The prevalence of suicidal behaviour in the general population ranged from 5% (39) to 75% (37). In the primary care setting without a pre-test, the prevalence ranged from 1%–3% (38, 40). In the primary care setting with a pre-test for depression, the prevalence of
suicidal behaviour was 35% (41).

The sensitivity of the different index tests varied between 26% (39) and 100% (38), the specificity between 64% (36) and 99% (38, 39). The PPV also showed a wide range from 6% (38) to 91% (37). The NPV had values from 83% (41) to 100% (38).

The certainty of the evidence was high in two studies (38, 39), moderate in three studies (37, 40, 41) and low in one study (36). Results of the individual studies are presented in Table 4. The measures of diagnostic accuracy sensitivity and specificity are visualized in a forest plot in Figure 2.

**Target condition suicidal ideation**

Chamberlain et al. (2009) evaluated the K10 as a screening instrument for suicidal ideation. We calculated sensitivity and specificity with 95% confidence intervals for the cut off value high (22) and very high (30). The calculated sensitivity was 26% and the specificity 99% for the very high cut off and 58% sensitivity and 93% specificity for the high cut off. The certainty of evidence was rated as high.

Millner et al. (2015) examined the Gate question suicide ideation as a single item index test against Follow up questions for suicidal ideation. After consulting the author, we chose Maybe I should kill myself as a cut off value for the reference standard and calculated a sensitivity of 96% and specificity of 72%. The certainty of evidence was moderate.

Cooper-Patrick et al. (1994) developed the SIS-Q, evaluating it against four suicidal behaviour questions of the Diagnostic Interview Schedule (DIS). The author reported a sensitivity of 84% with application of a cut off value of one related to the baseline Interview. The calculated specificity was 75%. The certainty of evidence was estimated as moderate.

Olfson et al. (1996) reported sensitivity of the single item index test Feeling suicidal with
86%. Amended specificity was 99%. The reference standard was a structured interview including questions on current and past suicidal behaviour. For the present review, the certainty of evidence was rated as high.

Target condition suicide plan

Van Spijker et al. (2014) developed the SIDAS and compared it to suicide plan items of the Columbia-Suicide Severity Rating Scale (C-SSRS). The author reported diagnostic accuracy measures for a cut off value of 1 with sensitivity of 86%, specificity of 67% and a cut off value of 21 with 40% sensitivity and 96% specificity. For the present review, the certainty of evidence was rated as low.

Olfson et al. (1996) evaluated the Feeling suicidal single item index test also against the reference standard In the last month, did you have a plan to kill yourself, a question from the structured interview. This resulted in a sensitivity of 83% and specificity of 98%. Furthermore, the author evaluated the single item index test Thoughts of death, sensitivity 100%, specificity 81% as well as the single item index test Wishing you were dead with a sensitivity of 92% and specificity of 93% in detecting patients with plans to commit suicide. For the present review, the certainty of evidence was rated as high.

Target condition suicide attempt

Millner et al. (2015) evaluated the Gate questions suicide attempt against a coding procedure for suicide attempt. Sensitivity was 89% and specificity 94%. For the present review, the certainty of evidence was rated as moderate.

Van Spijker et al. (2014) evaluated the SIDAS against the suicide attempt items of the C-SSRS. The author reported diagnostic accuracy measures for the cut off value of 1 with sensitivity of 84% and specificity of 64% and cut off value of 21 with sensitivity of 50% and specificity of 95%. For the present review, the certainty of evidence was rated as low.

Target condition suicidal behaviour
Uebelacker et al. (2011) examined the ability of the PHQ-9 - item 9 in detecting suicidal behaviour compared to the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) mood module. Reported sensitivity was 69% and specificity 84%. For the present review, the certainty of evidence was classified as moderate.

Discussion

We identified a total of 14 studies that assessed suicidal behaviour in the primary care setting or the general population. Of those, only six measured the required diagnostic accuracy. This review identified two further studies in the adult primary care population (40, 41) in addition to the already in the review of U.S. Preventive Services Task Force (24) described study by Olfson et al. (1996) (38) and three studies in the general population (36, 37, 39).

If pragmatically chosen benchmark values from Runeson et al. (2017) (50) with a required minimum level of sensitivity of 80% and specificity of 50% were applied, seven of nine index tests achieved these values: 1. SIS-Q, 2. Gate question suicide ideation, 3. Feeling suicidal 4. Thoughts of death, 5. Wishing you were dead, 6. SIDAS with a cut off value of 1, 7. Gate question for suicide attempt

All of the screeners, except of SIDAS, yielded in a high to moderate GRADEpro rating.

SIDAS was assessed as low certainty of evidence.

Six of the nine screening instruments were screening instruments with only one item.

Single item instruments are controversially discussed in the literature (37, 51–53). In the study by Millner et al. (2015), both single item index tests yielded acceptable sensitivities and specificities, but the author advises against using these questionnaires, since single item measures do not capture a large proportion of suicidal behaviour, and his statistical simulations showed that these lead to more erroneously classifications than questionnaires with more items (37). In contrast, item 9 of PHQ–9 was investigated by
Simon et al. (2013); the author reported that the response to item 9 could identify patients with an increased risk of suicide attempt and suicide (53).

Screening of patients at risk is recommended by several medical societies (18, 54, 55). In this context, it should also be considered that a sample of primary care patients at risk (e.g. patients with depression) and a sample of unselected primary care patients would have different prevalence of suicidal behaviour. The prevalence of a disease directly influences the values PPV and NPV (56). PPV and NPV are particularly interesting for physicians because they describe the probability that a patient with a positive (negative) test result is ill (not ill) (56). The populations of primary care patients and general population are usually settings with a low prevalence of suicidal behaviour, e.g. in the range from 0.01 (38) and 0.05 (39). The resulting NPV of around 95% underlines suicidal behaviour is a very rare condition with a pre-test probability of being healthy from 0.95 to 0.99. The resulting PPV of around 15–30% means at the same time that 70–85% are diagnosed as false positive. Therefore, as any other symptom/disease in the low prevalence range, prediction is limited (57). Thus it is necessary to identify potential patients at risk (e.g. depression) before screening for suicidal behaviour. Solely the study by Uebelacker et al. (2011) (41) investigated patients at risk (patients with depression) with regard to their suicidal behaviour.

Suicidal behaviour is described as a continuous process. It begins with occasional suicidal thoughts and can lead to the complete suicide (1). In primary care an early detection could appear to be useful. Suicidal patients would be detected at an early stage and further escalation to a suicide plan or suicide attempt could be prevented by direct intervention (3). With regard to the target condition of suicidal thoughts or suicidal behaviour (including suicidal thoughts) the K10, SIS-Q, Gate question suicide ideation and the PHQ-9 - item 9 were identified.
Furthermore, it is discussed in literature whether suicidality should be addressed directly. Chamberlain et al. (2009) (39) and Cooper-Patrick et al. (1996) (40) point out as an advantage that the K10 and SIS-Q do not directly confront patients with suicidal behaviour. However, national disease management guidelines for depression recommend direct enquiries (18, 58). Concerns about inducing suicidal behaviour through screening could be clarified (59).

A further approach is that hope and optimism contribute to mental health (3, 60, 61), protective factors should be included in the medical history (18, 48). Chang et al. (2013) found out that optimism is a protective factor for suicide risk in primary care patients (62). To our knowledge, there is only one short screener, the P4 (48), that assesses protective factors; however, the P4 was excluded from our analysis due to missing diagnostic accuracy measures.

A closer look at the PHQ–9 - item 9 is worthwhile. Although item 9 did not achieve the required sensitivity of 80% (cf. 69%), it has the advantage that it is already integrated into everyday practice as a depression screener (41). A prospective study in primary care and mental health care showed that the answer to item 9 had a strong predictive value for suicide attempts and suicide (53).

The choice of the reference standard plays a decisive role, as the results of the index test depend directly on the assessment of the reference standard. A diagnostic interview e.g. SCID-I (63) is supposed to be the highest standard. In the literature, however, different screening instruments have been referred to as the gold standard, including short questionnaires such as the C-SSRS (47, 64, 65). An agreement on a gold standard would be an important step for future research. This would improve the comparability of the results.

The NPV of the screening instruments was over 80%. If further studies could replicate
these results and confirm them in a setting with patients at risk, the screening instruments could help the GP to rule out suicidal behaviour.

Limitations

The studies assessed different index tests. All index tests were examined in only one study, which limits the interpretability of the results. Moreover, the studies included reported different reference standards. As there is no gold standard in the assessment of suicidality, a comparison of the diagnostic accuracy values of the index tests is difficult. Besides overall risk of bias and applicability concerns were moderate. This could limit the interpretability of the results.

When evaluating with GRADE, it must be critically considered that each index test was only evaluated in a single study.

As only studies measuring diagnostic test accuracy were included in our review-describing criterion validity (66) - articles examining other criteria of psychological tests (67) were not evaluated further (42–49). For promising short questionnaires like the SBQ-R, P4 or the CSSRS we did not identify studies in the primary care setting or the general population which measured their diagnostic accuracy.

In any systematic review there remains a risk that articles were not identified because of the selection of search terms or the erroneous exclusion in title and abstract screening.

The review was limited to articles in English and German language. Self-reported suicidal behaviour may be limited by individuals’ motivation to reveal their suicidal plans or intent; this affects the results of the studies (68).

Conclusion

Screening of unselected primary care patients for suicidal behaviour is little effective. In contrast screening of patients at risk is an important step in diagnostics of suicidal
behaviour (69). In Germany GPs have competence in this field due to their specialized training (70, 71). Promising short questionnaires such as PHQ–9 - item 9, or P4 should be included in future research in order to assist GPs in their judgement of suicidal behaviour of patients at risk. However, the final assessment is always based on the clinical judgement of the attending physician (72). The agreement on a gold standard would be an important step to enable comparison of results of different index tests. In case of a positive response, further exploration of the patients would be necessary.

List Of Abbreviations

ACSS Acquired Capability for Suicide Scale
BSS Beck Scale for Suicide Ideation
C-SSRS Columbia-Suicide Severity Rating Scale
DIS Diagnostic Interview Schedule
GP general practitioner
JG Jochen Gensichen
K10 Kessler Psychological Distress Scale
KH Kathrin Halfter
KL Karoline Lukaschek
MF Milena Frank
NPV negative predictive value
PHQ–9 Patient Health Questionnaire–9
PPV positive predictive value
QUADAS–2 Quality Assessment of Diagnostic Accuracy Studies - version 2
RASS Risk Assessment Suicidality Scale
SABCS Suicidal Affect-Behaviour-Cognition Scale
SBQ-R Suicide Behaviours Questionnaire Revised
SCID-I Structured Clinical Interview for DSM-IV Axis I Disorders
SIDAS Suicidal Ideation Attributes Scale
SIS-Q Suicidal Ideation Screening Questionnaire

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Availability of data and material
The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests

Funding
There was no funding of the review.

Authors’ contributions
KL and MF wrote the manuscript, performed the title and abstract screening, the full text analysis, the data collection, the risk of bias assessment and rated the certainty of evidence. MF and KH performed the statistical analysis. JG and AS revised the manuscript critically for clinical content.
All authors have read and approved the manuscript.

Acknowledgements
The authors would like to thank librarian Dr Andreas Will (LMU University Library) for supporting the development of the electronic search and data management.

References

1. 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.

2. Bertolote JM, Fleischmann A, De Leo D, Bolhari J, Botega N, De Silva D, et al. Suicide attempts, plans, and ideation in culturally diverse sites: the WHO SUPRE-MISS community survey. Psychol Med. 2005;35(10):1457–65.

3. O’Connor RC, Nock MK. The psychology of suicidal behaviour. Lancet Psychiatry. 2014;1(1):73–85.

4. 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.

5. 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.

6. 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). European psychiatry. 2015;30(7):874–9.

7. Ladwig KH, Klupsch D, Ruf E, Meisinger C, Baumert J, Erazo N, et al. Sex- and age-related increase in prevalence rates of death wishes and suicidal ideation in the community: results from the KORA-F3 Augsburg Study with 3,154 men and women, 35 to 84 years of age. Psychiatry research. 2008;161(2):248–52.

8. Forkmann T, Brahler E, Gauggel S, Glaesmer H. Prevalence of suicidal ideation and
related risk factors in the German general population. J Nerv Ment Dis. 2012;200(5):401–5.

9. 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. The British journal of psychiatry: the journal of mental science. 2008;192(2):98–105.

10. Schulberg HC, Bruce ML, Lee PW, Williams Jr JW, Dietrich AJ. Preventing suicide in primary care patients: the primary care physician’s role. General hospital psychiatry. 2004;26(5):337–45.

11. 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.

12. 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.

13. Turecki G, Brent DA. Suicide and suicidal behaviour. Lancet. 2016;387(10024):1227–39.

14. Horwitz AG, Czyz EK, King CA. Predicting future suicide attempts among adolescent and emerging adult psychiatric emergency patients. Journal of Clinical Child & Adolescent Psychology. 2015;44(5):751–61.

15. Brown GK, Beck AT, Steer RA, Grisham JR. Risk factors for suicide in psychiatric outpatients: a 20-year prospective study. Journal of consulting and clinical psychology. 2000;68(3):371.

16. 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. European Psychiatry. 2010;25(7):427–31.

17. World Health Organization. Preventing suicide: A global imperative.[cited 09.07.2019]. Available from: https://apps.who.int/iris/bitstream/handle/10665/131056/9789241564779_eng.pdf?
sequence = 1.

18. DGPPN, BÄK, KBV, AWMF. S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression—Kurzfassung [cited 28.02.2019]. Available from: www.depression.versorgungsleitlinien.de.

19. 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.

20. 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. Scandinavian journal of public health. 2019;47(1):9–17.

21. 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. The British journal of general practice: the journal of the Royal College of General Practitioners. 2009;59(568):825–32.

22. 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-.

23. 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.

24. 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. Annals of internal medicine. 2013;158(10):741–54.

25. Brähler E, Zenger M, Kemper CJ. Psychologische und sozialwissenschaftliche Kurzskalen: Standardisierte Erhebungsinstrumente für Wissenschaft und Praxis. Berlin: MWV; 2015.

26. 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.

27.Wolters Kluwer. Ovid.[cited 20.08.2019]. Available from: http://ovidsp.dc2.ovid.com.emedien.ub.uni-muenchen.de.

28.Cochrane Collaboration. Cochrane Library.[cited 13.06.2019]. Available from: https://www-cochranelibrary-com.emedien.ub.uni-muenchen.de/.

29.EBSCO Information Services GmbH. EBSCO.[cited 13.06.2019]. Available from: http://web.a.ebscohost.com.emedien.ub.uni-muenchen.de/.

30.Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic reviews. 2016;5(1):210.

31.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. Clinical chemistry. 2003;49(1):7-18.

32.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. Annals of internal medicine. 2011;155(8):529-36.

33.McMaster University. GRADEpro Guideline Development Tool [Software].[cited 20.08.2019]. Available from: gradepro.org.

34.Schünemann HJ. Introductory courses for GRADE and Summary of Findings Tables - Inconsistency.[cited 13.06.2019]. Available from: https://cebgrade.mcmaster.ca/Inconsistency/index.html.

35.Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine. 2009;151(4):264-9.

36.Van Spijker BA, Batterham PJ, Calear 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 and Life-Threatening Behavior. 2014;44(4):408–19.

37. Millner AJ, Lee MD, Nock MK. Single-item measurement of suicidal behaviors: Validity and consequences of misclassification. PLoS ONE. 2015;10(10):e0141606.

38. Olfson M, Weissman MM, Leon AC, Sheehan DV, Farber L. Suicidal ideation in primary care. J Gen Intern Med. 1996;11(8):447–53.

39. 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.

40. Cooper-Patrick L, Crum RM, Ford DE. Identifying suicidal ideation in general medical patients. Jama. 1994;272(22):1757–62.

41. 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. The primary care companion to CNS disorders. 2011;13(1).

42. 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 and Life-Threatening Behavior. 2015;45(4):448–60.

43. 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.

44. 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. Journal of affective disorders. 2019;245:779–87.

45. 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.
Journal of affective disorders. 2012;138(3):449–57.

46. 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. Journal of Psychopathology and Behavioral Assessment. 2017;39(2):291–302.

47. 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.

48. 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.

49. 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). PPmP-Psychotherapie· Psychosomatik· Medizinische Psychologie. 2018;68(08):346–52.

50. 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.

51. Hom MA, Joiner Jr TE, Bernert RA. Limitations of a single-item assessment of suicide attempt history: Implications for standardized suicide risk assessment. Psychological Assessment. 2016;28(8):1026.

52. Gardner DG, Cummings LL, Dunham RB, Pierce JL. Single-item versus multiple-item measurement scales: An empirical comparison. Educational and psychological measurement. 1998;58(6):898–915.

53. 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.

54. 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.

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

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

57. 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.

58. National Institute for Clinical Excellence. Depression: management of depression in primary and secondary care: National Institute for Clinical Excellence; 2004. Available from: http://www.scamfyc.org/documentos/depresion%20NICE.pdf.

59. 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.

60. Gallagher MW, Lopez SJ. Positive expectancies and mental health: Identifying the unique contributions of hope and optimism. The Journal of Positive Psychology. 2009;4(6):548-56.

61. Teismann T, Brailovskaia J, Margraf J. Positive mental health, positive affect and suicide ideation. Int J Clin Health Psychol. 2019;19(2):165-9.
62. 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? Cognitive Therapy and Research. 2013;37(4):796-804.

63. 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.

64. 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.

65. 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 [cited 20.08.2019]. 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.

66. Rutjes A, Reitsma J, Coomarasamy A, Khan K, Bossuyt P. Evaluation of diagnostic tests when there is no gold standard. A review of methods. Health Technology Assessment-Southampton. 2007;11(50).

67. Glaesmer H, Forkmann T, Dinkel A, Wahl I, Sattel H, Huber D, et al. Kriterienkatalog zur Beurteilung psychodiagnostischer Selbstbeurteilungsinstrumente-Empfehlung des Deutschen Kollegiums für Psychosomatische Medizin (DKPM). PPmP-Psychotherapie Psychosomatik-Medizinische Psychologie. 2015;65(07):246-54.

68. Glenn CR, Cha CB, Kleiman EM, Nock MK. Understanding suicide risk within the Research Domain Criteria (RDoC) framework: insights, challenges, and future research considerations. Clinical psychological science. 2017;5(3):568-92.

69. Forkmann T, Teismann T, Glaesmer H. Diagnostik von Suizidalität. Göttingen: Hogrefe
70. Kruse W, Saß H. Curriculum Psychosomatische Grundversorgung-Basisdiagnostik und Basisversorgung bei Patienten mit psychischen und psychosomatischen Störungen einschließlich Aspekte der Qualitätssicherung. Köln: Bundesärztekammer 2001. Available from: https://www.aerzte-seminar.de/wp-content/uploads/2018/07/Curriculum_PsychosomatischeGrundversorgung_BAEK.pdf.

71. DEGAM. Kompetenzbasiertes Curriculum Allgemeinmedizin. [cited 19.07.2019]. Available from: https://www.weiterbildung-allgemeinmedizin.de/downloads/Curriculum_01-10-15.pdf.

72. Teismann T, Dorrmann W. Suizidalität. Göttingen: Hogrefe Verlag; 2014.

Tables

Table 1 Key characteristics included studies

| Setting               | Target condition     | Index test | Reference standard                        |
|-----------------------|----------------------|------------|------------------------------------------|
| Telephone interview   | Suicidal ideation   | K10        | 4 questions subscale C in the General Health Questionnaire 28 |
| Interview             |                      | 10 items   | Interview-administered                   |
| Telephone             |                      | Interview- |                                          |
| interview             |                      | self-administered |                                          |
sub-Saharanans with telephone connection, randomly selected by the South Australian Monitor.
| Online survey | Suicidal ideation | Suicide attempt |
|---------------|-------------------|-----------------|
|               | 1. Gate question suicide attempt | 1. Action coded as suicide attempt |
|               | Single item Self-report | 2. Follow up questions for suicide ideation |
|               | 2. Gate question suicide ideation | Single item Self-report |
| Online survey | Suicide plan   | SIDAS     | Part of the C-SSRS |
|--------------|----------------|-----------|--------------------|
|              | Suicide        | 5 items   |                    |
|              | attempt/preparation | Self-report |                    |
Interviewed as part of the National Institute of Mental Health Epidemiologic Catchment Area surveys.

Suicidal ideation was assessed using the SIS-Q, a 4-item interview-administered questionnaire of the DIS.

Four suicidal behaviour questions of the DIS.
| Primary care clinic | Suicidal ideation | 1. Feeling suicidal |
|---------------------|-------------------|-------------------|
|                     |                   | Single item Self-report |
|                     |                   | 2. Wishing you were dead |
|                     |                   | Single item Self-report |
|                     |                   | 3. Thoughts of death |
|                     |                   | Single item Self-report |
|                     |                   | 4. Feeling suicidal |
|                     |                   | Single item Self-report |

1-3. Question of the structured interview from the research nurse: „In the last month, have you had a plan to kill yourself?“

4. Structured interview research nurse that includes questions on current and past suicidal behaviour.

Suicidal behaviour

PHQ-9 - item 9

Single item

Self-report

Suicide item on the mood module of the SCID-I
Table 2 Studies not assessed further because of missing diagnostic accuracy measurements
### Table 4 Results of individual studies

| First author; Year; Country          | Test                        | Number of items |
|-------------------------------------|-----------------------------|-----------------|
| Batterham, P. J.; 2015; Australia   | Interpersonal-Psychological Theory of Suicidal Behaviour | 12 items        |
| Dube, P.; 2010; USA                 | P4                          | 4 items         |
| Fountoulakis, K.; 2012; Greece      | RASS                        | 12 items        |
| Glaesmer, H.; 2018; Germany         | SBQ-R                       | 4 items         |
| Harris, K. M.; Study 2 and 3; 2015; Australia | SABCS            | 6 items         |
| Kliem, S.; 2017; Germany            | BSS-Screen                  | 5 items         |
| Turnell, A. I.; 2019; Australia     | Self-Hate Scale             | 7 items         |
| Rimkeviciene, J.; Study 2; 2017; Australia, Singapore | ACSS | 5, 7, 8, (20) items |

Due to technical limitations, Table 3 is only available as a download in the supplemental files section.
TP = True positive, FP = False positive, TN = True negative, FN = False negative,

| Author                          | Index test             | Reference standard | n  | T   | P   | F   | TN | F   | Sensitivity (95% CI) | Specificity (95% CI) | PPV (95% CI) | NPV (95% CI) | Certainty of evidence |
|---------------------------------|------------------------|--------------------|----|-----|-----|-----|-----|-----|----------------------|----------------------|-------------|-------------|----------------------|
| Chamberlain et al.              | K10 cut off 30         | GHQ-28 subscale    | 1  | 2   | 0   | 4   | 5   | 0   | 0.26 (0.23-0.30)     | 0.99 (0.98-0.99)     | 0.52 (0.46-0.58) | 0.96 (0.96-0.96) | high               |
| Chamberlain et al.              | K10 cut off 22         | GHQ-28 subscale    | 1  | 2   | 0   | 4   | 5   | 0   | 0.58 (0.54-0.62)     | 0.93 (0.92-0.93)     | 0.30 (0.28-0.33) | 0.98 (0.97-0.98) | high               |
| Millner et al.                  | Gate question suicide attempt | Action coded as suicide attempt | 1  | 2   | 0   | 4   | 5   | 0   | 0.89 (0.86-0.92)     | 0.94 (0.92-0.95)     | 0.89 (0.86-0.92) | 0.94 (0.92-0.95) | moderate           |
| Millner et al.                  | Gate question suicide ideation | Follow up question for suicide ideation | 1  | 2   | 0   | 4   | 5   | 0   | 0.96 (0.95-0.97)     | 0.72 (0.68-0.77)     | 0.91 (0.89-0.93) | 0.86 (0.82-0.90) | moderate           |
| van Spijker et al.              | SIDAS cut off 1        | C-SSRS-suicide plan | 1  | 2   | 0   | 4   | 5   | 0   | 0.16 (0.18-0.20)     | 0.86 (0.80-0.90)     | 0.67 (0.64-0.70) | 0.34 (0.30-0.38) | low                |
| van Spijker et al.              | SIDAS cut off 21       | C-SSRS-suicide plan | 1  | 2   | 0   | 4   | 5   | 0   | 0.11 (0.12-0.13)     | 0.84 (0.77-0.89)     | 0.64 (0.61-0.66) | 0.22 (0.18-0.25) | low                |
| Cooper-Patrick et al.           | SIS-Q cut off 1        | DIS-suicidal behaviu r questions | 1  | 2   | 0   | 4   | 5   | 0   | 0.03 (0.03-0.05)     | 0.84 (0.77-0.90)     | 0.75 (0.74-0.76) | 0.08 (0.07-0.10) | low                |
| Ofsson et al.                   | Feeling suicidal - suicidal ideation | Structure d interview | 1  | 2   | 0   | 4   | 5   | 0   | 0.03 (0.03-0.05)     | 0.86 (0.67-0.95)     | 0.99 (0.98-0.99) | 0.76 (0.57-0.88) | high               |
| Ofsson et al.                   | Feeling suicidal - suicidal plan | Structure d interview | 1  | 2   | 0   | 4   | 5   | 0   | 0.01 (0.01-0.02)     | 0.83 (0.62-0.67)     | 0.98 (0.97-0.99) | 0.30 (0.15-0.46) | high               |
| Ofsson et al.                   | Thoughts of death      | Structure d interview | 1  | 2   | 0   | 4   | 5   | 0   | 0.01 (0.01-0.02)     | 1.00 (-)             | 0.81 (0.79-0.84) | 0.06 (0.03-0.09) | high               |
| Ofsson et al.                   | Wishing you were dead  | Structure d interview | 1  | 2   | 0   | 4   | 5   | 0   | 0.01 (0.01-0.02)     | 0.92 (0.76-0.78)     | 0.93 (0.92-0.95) | 0.14 (0.06-0.22) | high               |
| Uebelacker et al.               | PHQ-9 - item 9         | SCID-I             | 1  | 2   | 0   | 4   | 5   | 0   | 0.69 (0.55-0.80)     | 0.84 (0.76-0.90)     | 0.70 (0.56-0.81) | 0.83 (0.75-0.90) | moderate           |
Figures

Figure 1

PRISMA Flow Diagram (35)
Sensitivity and Specificity in the general population

Van Spijker et al.  SIDAS Cutoff 21 - suicide attempt
Van Spijker et al.  SIDAS Cutoff 21 - suicide plan
Van Spijker et al.  SIDAS Cutoff 1 - suicide attempt
Van Spijker et al.  SIDAS Cutoff 1 - suicide plan
Millner et al.  Gate question suicide ideation
Millner et al.  Gate question suicide attempt
Chamberlain et al.  K10 Cutoff 22
Chamberlain et al.  K10 Cutoff 30

Sensitivity and Specificity in the primary care population

Uebelacker et al.  PHQ-9, Item 9
Olsson et al.  Wishing you were dead
Olsson et al.  Thoughts of death
Olsson et al.  Feeling suicidal - suicide plan
Olsson et al.  Feeling suicidal - suicide ideation
Olsson et al.  SIS-Q Cutoff 1

Figure 2
Forest plot of sensitivity and specificity of index tests of included studies

Supplementary Files

This is a list of supplementary files associated with the primary manuscript. Click to download.
Additional file.docx
Additional file - PRISMA Checklist.docx
Table 3.jpg