



BMJ Open Development and psychometric evaluation of a questionnaire for the assessment of depression in primary care: a cross-sectional study

Clara Teusen ¹, Markus Bühner,² Alexander Hapfelmeier,^{1,3} Victoria von Schrottenberg,¹ Klaus Linde,¹ Jochen Gensichen,^{4,5} Antonius Schneider ¹

To cite: Teusen C, Bühner M, Hapfelmeier A, *et al.* Development and psychometric evaluation of a questionnaire for the assessment of depression in primary care: a cross-sectional study. *BMJ Open* 2024;**14**:e084102. doi:10.1136/bmjopen-2024-084102

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2024-084102>).

Received 09 January 2024
Accepted 27 June 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Clara Teusen;
clara.teusen@mri.tum.de

ABSTRACT

Objectives To develop a new questionnaire for the diagnostic assessment of depression adapted to the primary care setting by combining psychiatric criteria and heuristics of general practitioners (GPs). Psychometric evaluation of the new questionnaire and first validity evidence.

Design The questionnaire was developed using cognitive interviews with think-aloud technique. Factorial validity was then examined in a cross-sectional study.

Setting Primary care. Five general practices in Bavaria, Germany.

Participants 15 GPs, 4 psychiatrists/psychotherapists and 13 patients participated in cognitive expert interviews. A primary care sample of N=277 consecutive patients participated in the cross-sectional study.

Methods After consultation with experts and literature research, the questionnaire contained a self-rating part for patients and an external part for GPs. Items were then iteratively optimised using cognitive interviews. Factorial validity was examined. To estimate internal consistency, Cronbach's α was calculated. Validity was assessed by correlating the new questionnaire and the Patient Health Questionnaire-9 (PHQ-9).

Results The preliminary version of the two-part 'Questionnaire for the assessment of DEpression SYmptoms in Primary Care' (DESY-PC) comprised 52 items for patients (DESY-PAT-1: questions about patient's environment; DESY-PAT-2: questions about depression-specific symptoms) and 21 items for GPs (DESY-GP). The analysis of the DESY-PAT-1 revealed a one-factor solution ('environmental factors') with Cronbach's α of 0.55. The items of the DESY-PAT-2 were assigned to three factors, 'depressive cognitions', 'suicidality' and 'symptoms of fatigue', with Cronbach's α of 0.86, 0.79 and 0.85, respectively. Factorial analysis revealed two factors for the DESY-GP: 'depression symptoms' and 'medical history/external factors'. Cronbach's α was 0.90 and 0.59, respectively. After factorial analysis, the DESY-PAT was reduced to 28 items, and the DESY-GP was reduced to 15 items. Correlations of the DESY-PC with the PHQ-9 were high and significant, indicating convergent validity.

Conclusions The new questionnaire represents an innovative extension of depression questionnaires and could be particularly suitable for general practices.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ The participation of 32 experts in the construction of the questionnaire ensured that general practitioner-specific heuristics and patient-related characteristics of the primary care setting were incorporated into the new questionnaire.
- ⇒ Unlike other validated depression questionnaires, the new questionnaire includes not only psychiatric criteria for depression but also contextual factors relevant to general practice that may improve the diagnosis of depression.
- ⇒ It was not tested whether the questionnaire for the assessment of DEpression SYmptoms in Primary Care identifies depression more accurately than commonly used depression questionnaires, as we did not apply a Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders to confirm or rule out a depression diagnosis.

INTRODUCTION

The general practitioner (GP) is usually the first healthcare provider that patients consult.^{1–3} In most cases, GPs are also the gatekeepers for further diagnostics and treatment of patients with depression.^{4,5} However, identifying depression in primary care can be challenging when only somatic symptoms are reported, and patients do not explicitly mention their depressed mood.⁶ In addition to this challenge, the diagnosis of depression in primary care is further complicated by multimorbidity. Somatic complaints often overlap and mask symptoms of depression, so it can be difficult to distinguish between somatic disorders and depression.^{7,8} In any case, the initial diagnosis is essential for subsequent treatment.^{5,9} Thus, it is crucial that GPs follow a guideline-oriented diagnostic process and treatment, as the majority

of patients with depression are only seen in general practice.^{9 10}

In this context, it is important to note that depression is one of the most prevalent mental disorders.^{11–13} Various studies have reported a lifetime prevalence of depressive disorders ranging from 12% to 19%.^{12 14–16} Depression has a major impact on the lives of those who are affected, on their family members, and on their immediate environment. Therefore, it represents a considerable health problem for our society.^{17 18} Between 2005 and 2015, depression rose from the fourth to the third-leading cause of disability.¹⁹ Moreover, the WHO predicts that depression will be the largest burden of disease worldwide by 2030.²⁰ Hence, it is particularly important to improve the diagnosis and care of patients with depression and to optimise treatment processes.²¹ It is crucial to identify and treat people with depression in the early stages of their illness to prevent chronicity.²² In addition, proactive management of subthreshold depression can also protect affected individuals from developing major depression.²³

Standardised screening questionnaires could be one approach to improve the diagnosis of depression in primary care. However, expert panels such as the Canadian Task Force on Preventive Health Care do not recommend routine screening for depression in general practice.²⁴ Similarly, guidelines such as the UK National Institute for Health and Care Excellence guideline or the German National Disease Management Guideline (NDMG; in German: Nationale Versorgungsleitlinie, NVL) on Unipolar Depression do not explicitly call for routine screening. Nevertheless, both recommend it if risk factors for depression are present and the GP suspects depression.^{25 26} Although the Patient Health Questionnaire-9 (PHQ-9) has good sensitivity and specificity, previous studies have shown that screening for depression in primary care can result in a high rate of false-positives,^{27–32} leading to the misclassification of healthy patients as depressed. In addition, screening for depression has not been shown to improve mental health.³³ An alternative to screening in primary care could be the use of diagnostic tools as an aid to diagnosis if the clinician already suspects depression.

Furthermore, it was shown that standard diagnostic systems (e.g. International Statistical Classification of Diseases and Related Health Problems 10, ICD-10) do not work adequately in the GP context.^{22 34 35} GPs use their heuristics and rely on factors other than ICD-10/11 or DSM-V (Diagnostic and Statistical Manual of Mental Disorders V) criteria.^{10 36 37} The GP's intuition, the consideration of biopsychosocial factors and their impression during the watchful waiting process, especially when depression is suspected, could represent such heuristics.^{35 38} While several studies have highlighted the impact of heuristics on medical decision-making,^{7 39} current questionnaires for depression do not incorporate the GP perspective so far.^{35 38} Considering GP heuristics and their perspective alongside the inclusion of psychiatric criteria could improve diagnostic decision-making and might be

superior for diagnosing depression in the primary care setting.⁴⁰ To our knowledge, no such questionnaire is adapted to the primary care setting and considers GP heuristics, thought processes and criteria for measuring depression. Therefore, a questionnaire that measures both psychiatric criteria or typical symptoms of depression and GP heuristics should be introduced in general practice. The planned questionnaire is, therefore, not intended as a classic screener but primarily as a diagnostic aid in general practice for patients who are considered to be at increased risk of depression.

In this article, we describe: (1) The development of a new questionnaire for the assessment of depression adapted to the GP setting, which considers GP heuristics and psychiatric criteria. (2) The psychometric evaluation of the new questionnaire and the first validity evidence in a primary care sample of N=277 patients.

METHODS

Development of the preliminary questionnaire

The first draft of the questionnaire was based on practical considerations, the clinical experience of the research team and the consideration of the main depression criteria from ICD-10. An initial literature review and discussions with three experienced GPs helped to refine the wording and number of items used. The first draft of the questionnaire was further developed by conceptual considerations of questionnaire construction and the consideration of commonly used screening questionnaires for depression, which were found to be relevant in a thorough literature review.^{30 41–45}

In the next step, the questionnaire design and content were iteratively optimised through cognitive expert interviews with GPs, psychiatrists/psychotherapists and patients. During the cognitive interviews, participants had to complete the new questionnaire by thinking out loud. We used this technique to detect inconsistencies, missing information/items or information about items that were difficult to understand. The cognitive think-aloud technique is optimal for capturing thought processes.⁴⁶ The idea was to consider psychiatric criteria and aspects essential to the GPs and their patients. The interviews were audiotaped and continuously analysed by the authors (CT, AS and MB), who discussed the plausibility of the suggestions and then iteratively incorporated them into the questionnaire before showing the revised version to the next interview partner. This process was conducted from April to October 2021 until construct saturation occurred, and no further far-reaching suggestions for improvement were made. GP interview partners were recruited through the Bavarian practice-based research network; patients were recruited through GP referral and recruitment on a psychiatric ward. Psychiatrists/psychotherapists were motivated to participate in an interview by direct invitation. The 32 interview partners gave written informed consent.

The development process resulted in a two-part questionnaire: a self-rating questionnaire for general practice patients and an external rating questionnaire for GPs. As a next step, a cross-sectional study was conducted, and the factorial structure of the new two-part questionnaire was examined to identify its factorial and psychometric properties.

Study design, procedure and participants during the evaluation of the questionnaire

The cross-sectional study was performed between March and July 2022 in five general practices in Bavaria, Germany. This study was registered with the German Clinical Trials Registry (DRKS-ID: DRKS00028950). Inclusion criteria were an age of at least 18 years, sufficient knowledge of the German language and a signed consent form. All patients were approached consecutively (i.e. without preselection) on certain days at regular intervals in the GP's waiting room, regardless of their reason for the encounter with the GP. As the new questionnaire was to be tested first, patients with and without depression had to fill it out in order to examine how well the questionnaire discriminated between these patients. After giving informed consent, they were asked to complete a self-report questionnaire consisting of our newly developed questionnaire and the PHQ-9. After the consultation with the patient, the GP had to fill in the external rating part of the newly developed questionnaire.

Instruments

Preliminary questionnaire for the assessment of DEpression SYmptoms in Primary Care

Our newly developed questionnaire for the assessment of DEpression SYmptoms in Primary Care (DESY-PC) contains a self-rating part for patients and an external rating part for GPs. As part of the following analysis of the factorial structure, the number of items in both questionnaire parts was reduced (see online supplemental material for the preliminary version of the DESY-PC). The questionnaire was originally written in German. To present an English version as part of this article, we translated the questionnaire back and forth between German and English using an online machine translation service (DeepL Translator, DeepL.com). The English version was then reviewed with a native speaker who is fluent in German.

Preliminary self-rating part for patients (DESY-PAT): This part contains 13 items with general questions about the patient's environment (DESY-PAT-1), followed by 29 questions about depression-specific symptoms (DESY-PAT-2). All items are presented in a closed-answer format (yes/no). This preliminary part is depicted in the online supplement (online supplemental material S1).

Preliminary external rating part for GPs (DESY-GP): This part examines the presence of depression in the patient from the GP's point of view. The questionnaire part comprises 21 items, which are presented in a

closed-answer format (yes/no). This preliminary part is depicted in the online supplemental material S2.

Patient Health Questionnaire-9

The validated questionnaire PHQ-9 is used to detect patients at high risk for depression.⁴⁷ The PHQ-9 is a module of the PHQ-D. It includes nine items and can be used to determine the severity of depression. A cut-off score of ≥ 10 is used to indicate a high risk of depression.⁴⁸ In this study, the PHQ-9 is used as a comparative questionnaire for the convergent validity of the newly developed DESY-PC.

Further recorded data

Demographic data were examined with respect to age, gender, origin, sociodemographic background and reason for encounter. Additionally, the permanent diagnoses noted in the GP's computer system, the current reason for the encounter noted by the GP and the medication were recorded.

Data analysis

Descriptive statistics of quantitative or qualitative data are mean (M), SD and range, or absolute and relative frequencies.

We conducted an explorative factor analysis to assess the factorial validity of the questionnaire scales, DESY-GP, DESY-PAT-1 and DESY-PAT-2. We used the maximum likelihood (ML) method of the R package 'psych' with polychoric correlations and continuity correction.⁴⁹ We applied an oblimin rotation because the occurring factors were assumed to be correlated. The criterion for factor extraction was based on the results of the parallel analysis (polychoric correlations with ML-estimation and 5000 iterations). Additionally, we used the Minimum Average Partial Test (MAP-Test) and a series of ML model tests to determine the number of factors. This method was also used for factor extraction since overfactoring is less severe than underfactoring.⁵⁰ Afterwards, confirmatory factor analysis using the R package 'lavaan'⁵¹ with weighted least square mean and variance adjusted (WLSMV) was applied to detect violations of local fit. The model fit was assessed with TLI (Tucker-Lewis Index) and RMSEA (root mean square error of approximation). 90% confidence intervals (90% CI) were calculated. For the item analysis and the associated item selection, the item statistics (mean, SD, skewness) and the intercorrelations of the items were determined.

To estimate the internal consistency, we calculated Cronbach's coefficient α (Cronbach's α) for each scale of the DESY-PC as a minimum estimate of reliability. The PHQ-9 was used for convergent validation, which was estimated by correlating the DESY-PC and the PHQ-9. The associations between the scales of DESY-GP, DESY-PAT-1, DESY-PAT-2 and PHQ-9 were assessed with Pearson correlation coefficients and respective correction for attenuation. Items within a factor were 0/1 dummy-coded and summed, and corresponding sum scores were used

to calculate Pearson correlation coefficients. We used SPSS V.26.0 (IBM) and R V.4.1.0 (The R Foundation for Statistical Computing, Vienna, Austria) for statistical analyses. Hypothesis testing was performed at exploratory 5% significance levels.

Patient and public involvement

During the development of the questionnaire, we consulted a patient representative from the POKAL (Predictors and Outcomes in primary depression care) study group advisory board (DFG-GRK 2621), who advised us on the presentation and wording of the questionnaire and its application. Their approval was obtained before the questionnaire was used in the cross-sectional study. In addition, we sought advice from 13 primary care and psychiatric patients during the iterative development of the questionnaire.

RESULTS

Development of the DESY-PC

The first draft of the DESY-PC contained a distinct questionnaire part for GPs (DESY-GP) and consisted of 10 items with a closed-answer format (yes/no). After the revision of three experienced GPs, two items were added to the questionnaire, the wording of the present items was slightly modified and the structure was adjusted. The following systematic literature review resulted in additional changes: the order of the items was changed to guide the GP through the questions in a reasonable sequence, and items about family history of mental illness and medication replaced items regarding obesity and sleep. Besides, after careful conceptual considerations, the DESY-PC was extended by a separate self-rating questionnaire part for primary care patients (DESY-PAT). This questionnaire part was based on common depression questionnaires and contained 34 items with a closed answer format (yes/no).

The questionnaire construction process was followed by the iterative optimisation of the two-part questionnaire during 32 cognitive interviews with 15 GPs, 4 psychiatrists/psychotherapists and 13 patients. The cognitive thinking aloud procedure revealed that some items and questions were formulated too vague or that other questions were still missing. As a result, the number of items of the DESY-GP increased from 12 to 21. The DESY-PAT was split into two sections and contained 13 items about the patient's environment and 29 items regarding depression-specific symptoms, respectively. Various recommendations were made to change the wording and to improve the comprehensibility. The corresponding adjustments were made to finalise the development process. During this iterative development process, construct saturation was reached after interviewing 32 experts when no additional comments came up. The preliminary version of the two-part DESY-PC comprised 21 items for GPs (DESY-GP) and 13 plus 39 items for patients (DESY-PAT-1/2)

with a closed answer format (yes/no) after the iterative construction process.

Results of the cross-sectional study

Sample characteristics

From March to July 2022, 458 primary care patients were consecutively contacted in the waiting rooms of 5 general practices with 12 GPs in Bavaria. 286 patients agreed to participate in the study, and 277 signed the consent form and completed the questionnaire that was handed out to them (see [figure 1](#)). The mean age of the participants was 53.7 years (SD=18.2 years), and 55.2% were female. 15.2% of patients showed PHQ-9 sum scores ≥ 10 . For further sociodemographic descriptions, see [table 1](#).

DESY-PC: factorial validity and assessing scale internal consistency

DESY-PAT: The analysis of the DESY-PAT-1 ([table 2](#)) included n=240 (of N=277) usable cases (cases with missing values were removed). Although the parallel analysis suggested one factor, the MAP-Test indicated a three-factor solution, and the ML-tests indicated eight factors. Thus, we conducted an exploratory factor analysis with eight factors since overfactoring is a less severe problem than underfactoring.⁵⁰ We decided to select from each factor the item with the highest loading to build a content valid short scale. The DESY-PAT-1 now comprised eight essential items that were assigned to one factor, which measures 'environmental factors'. The loadings, communality, mean, SD, factor loadings and skewness are presented in [table 2](#). We tested the model with a WLSMV confirmatory factor analysis. An RMSEA of 0.05 (90% CI 0.00 to 0.08) and TLI of 0.81 were found. For the DESY-PAT-1 scale, Cronbach's α was 0.55 ('environmental factors').

The analysis of the DESY-PAT-2 ([table 2](#)) included n=248 (of N=277) usable cases. Before we started the analysis, item 28 ('In the last 2 weeks, have you tried to compensate for unpleasant feelings by using other addictive substances (eg, cannabis, ecstasy, cocaine, pills)?') of the DESY-PAT-2 was removed because there was too little variance in the response behaviour of the patients (too many 'no' answers). Since the parallel analysis revealed only one factor, and the model tests were significant for each solution, we decided to use the MAP-Test to achieve a higher resolution of factors. The MAP-Test revealed a three-factor solution. We removed eight items to reduce redundancy and to obtain a short scale that was as content-valid as possible. The exclusion of the items was discussed with a team of experts and finally approved. Therefore, the final DESY-PAT-2 comprised 20 items that were assigned to 3 factors: Factor 1 measures 'depressive cognitions', using nine items; factor 2 measures 'suicidality', using five items and factor 3 measures 'symptoms of fatigue', using six items. The loadings, communality, mean, SD, factor loadings and skewness are presented in [table 2](#). We tested the model with a WLSMV confirmatory factor analysis. An RMSEA of 0.05 (90% CI 0.03 to 0.06) and TLI of 0.92 were found in the confirmatory factor analysis. For the

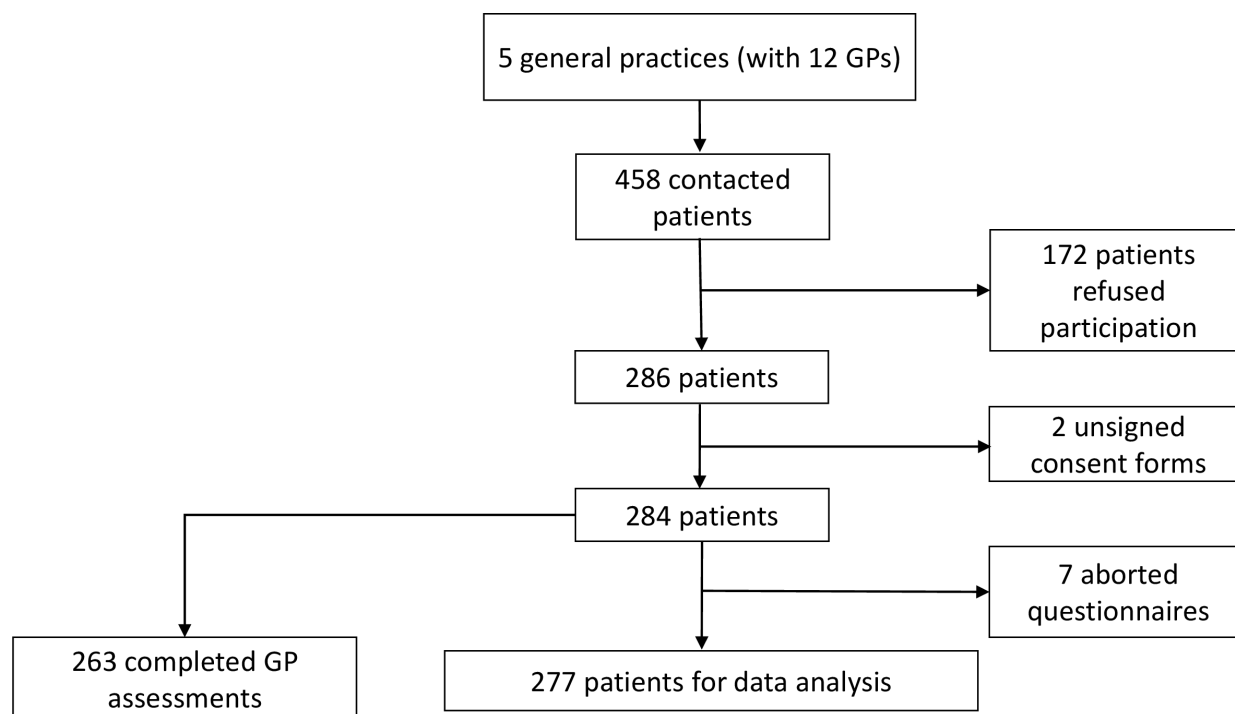


Figure 1 Flow chart of participants. GP, general practitioner.

DESY-PAT-2 scales, Cronbach's α was 0.86 ('depressive cognition'), 0.79 ('suicidality') and 0.85 ('symptoms of fatigue'). Additionally, we analysed the intercorrelations between the three DESY-PAT-2 scales, which ranged from 0.40 to 0.63. 'Depressive cognition' and 'suicidality' had the highest correlation ($r=0.63$), followed by 'depressive cognition' and 'symptoms of fatigue' ($r=0.51$). The lowest correlation was found between 'suicidality' and 'symptoms of fatigue' ($r=0.40$).

DESY-GP: For the factor analysis of the DESY-GP (table 3), we used the data of $n=263$ (of $N=277$) completed GP assessments. Before we started the analysis, item 20 ('For women: is a hormonal contraceptive being utilised?') of the DESY-GP was removed only for the analysis because this item produced, as expected, too many missing values. The item was also unable to capture any necessary additional information in terms of content and was, therefore, finally removed from the questionnaire. Although the parallel analysis suggested one factor, the MAP-Test indicated a two-factor solution, and a series of ML tests indicated eight factors. Thus, we conducted an exploratory factor analysis with eight factors. For factor 1, we selected six items out of seven representing 'depression symptoms'. One item (item 6, 'Is there evidence of increased fatigue and/or exhaustion?') was removed since there was a low loading on the main factor and similar high loadings on two other factors. The remaining factors consisted of only one or two items. We took the items with the highest loadings from these factors to build a content-valid factor, 'medical history/external factors', consisting of seven items. One item remained a universal item; even if this item did not load high enough on any factor, its requested content is considered necessary for

the questionnaire ('Have there ever been depressive phases?'). The loadings, communality, mean, SD, factor loadings and skewness are presented in table 3.

We tested both measurement models separately with a WLSMV confirmatory factor analysis. An RMSEA of 0.04 (90% CI 0.00 to 0.08) and TLI of 0.98 could be found in the confirmatory factor analysis for 'depressive cognitions'. For the factor 'medical history/external factors', an RMSEA of 0.04 (90% CI 0.00 to 0.08) and a TLI of 0.89 could be found. For the DESY-GP scales, Cronbach's α was 0.59 and 0.90 concerning 'medical history/external factors' and 'depression symptoms', respectively.

Convergent validity

The correlations of the DESY-PC and its subscales with the PHQ-9 all reach statistical significance. The correlation of the PHQ-9 with the DESY-PAT-1 and the DESY-PAT-2 is $r=0.57$ and $r=0.81$, respectively. In contrast to these high correlations, the DESY-GP only shows a moderate correlation of $r=0.45$ with the PHQ-9. Detailed correlations between DESY-PC and PHQ-9 can be found in figure 2. The distribution of observations is displayed by histograms and density plots on the diagonal. The lower triangle shows dot plots with a linear regression fit. The upper triangle shows Pearson correlation coefficients and a respective correction for attenuation.

DISCUSSION

The newly developed two-part questionnaire (DESY-PC) showed different factors for the self-rating part for patients (DESY-PAT) and for the external rating part for GPs (DESY-GP). The DESY-PAT consisted of two

Table 1 Characteristics of patients (N=277)

Variable (missing values)	Absolute frequency (percentage) or mean±SD (range)
Age in years (13)	53.7±18.2, (min.=18.1, max.=94.3)
Sex (1)	
Female	153 (55.2)
Diverse	8 (3.9)
Size of residence (27)	
<10 000 inhabitants	93 (33.6)
10 000–100 000 inhabitants	115 (41.5)
>100 000 inhabitants	42 (15.2)
Marital status (2)	
Married or in relationship	191 (69.0)
Divorced/widowed/single/other	79 (28.5)
Multiple answers	5 (1.8)
German nationality (27)	234 (84.5)
With children (7)	193 (69.7)
Highest level of general education completed (1)	
No secondary general school-leaving certificate	3 (1.1)
Secondary general/intermediate school-leaving certificate/ other/ multiple answers	172 (62.1)
High school diploma	101 (36.5)
Vocational qualification (4)	
No vocational training	5 (1.8)
Vocational qualification/other/ multiple answers	198 (71.6)
Higher education degree	70 (25.3)
Currently employed (9)	165 (59.6)
Diagnosis of depression detected in the past (5)	64 (23.1)
Present chronic disease(s) (0)	218 (78.7)
PHQ-9 ≥10 (4)	42 (15.2)
max., maximum; min., minimum; PHQ-9, Patient Health Questionnaire-9.	

parts. The DESY-PAT-1 presented a one-factor structure measuring ‘environmental factors’ for depression. During the development process of the questionnaire, the corresponding items in the DESY-PAT-1 were strongly influenced by the patients’ understanding of depression and by what they thought could play an essential role in the development of a depressive disorder. Therefore, the items of the DESY-PAT-1 go beyond validated depression questionnaires, like the PHQ-9, which primarily ask about commonly used psychiatric symptoms of depression, such as cognitive, emotional, physiological and behavioural

symptoms.⁴⁷ Although impairments in social, family and occupational functioning are also mentioned in the standard diagnostic criteria for depression,⁵² they have not yet been included in validated depression questionnaires.⁴⁵ The newly developed items in the DESY-PAT-1 focus on such environmental and contextual factors that can promote the onset of depression⁵³ and might play an essential role in diagnostic decision-making in general practice.³⁵ Environmental and contextual factors for depression can be very diverse and, when combined into a single factor, can lead to the relatively low internal consistency of 0.55 that we observed. The applicability of the DESY-PAT-1 requires further research to validate the findings and to demonstrate the diagnostic usefulness.

The DESY-PAT-2 showed a three-factor structure with one factor measuring ‘depressive cognitions’, another factor representing ‘suicidality’ and a third factor capturing ‘symptoms of fatigue’. The factor ‘depressive cognitions’ measures clinically relevant cognitive symptoms of depression, which are similarly captured, for example, by the PHQ-9.⁴⁷ The distinct factor ‘suicidality’ captures the proximity to death. This concept appears to be essential in the context of depression and should not be neglected during the process of diagnostic decision-making.⁵³ The concept of fatigue and lack of energy, captured by the third factor, is particularly striking and represents a crucial aspect during diagnostic decision-making of depression.⁵³ Many depressive primary care patients show reduced energy or fatigue symptoms, so this factor can be considered specific to the primary care setting.⁵⁴ The internal consistency of these three factors varied from 0.86 for ‘depressive cognition’, 0.79 for ‘suicidality’, to 0.85 for ‘symptoms of fatigue’. The results show that this part of the questionnaire measures three relevant aspects of depression in the primary care setting with sufficient precision to use the questionnaire for psychometric single-case diagnostic.

The items of the external rating part for GPs (DESY-GP) could be assigned to two independent factors, ‘depression symptoms’ and ‘medical history/external factors’. Besides, one universal item (‘Have there ever been depressive phases?’) was created. The internal consistency of the DESY-GP factors ranged from high, 0.90 for ‘depression symptoms’, to low, 0.59 for ‘medical history/external factors’. The first factor captures the symptoms of depression that GPs consider by comparing their impression of the patient in the current consultation with their experience of previous encounters with the same patient. In doing so, GPs take into account their in-depth knowledge of the patient, given by their shared medical history and familiarity, which ensures effective decision-making when considering standard psychiatric criteria for depression.^{54 55} However, the symptom count of standard diagnostic criteria should not be the only means for diagnosing depression in general practice. In addition, aetiological and contextual considerations are crucial for diagnostic decision-making.³⁵ Therefore, the DESY-GP also focuses on external factors of depression by

Table 2 ML-factor analysis with loadings of the DESY-PAT-1 and ML-factor analysis based on polychoric correlations with rotated loadings of the DESY-PAT-2, descriptive values

DESY-PAT-1	Factor		h^2	M	SD	r_{it}	V
Items	1 (environmental factors)*						
5 Do you currently have any financial difficulties?	0.86		0.74	0.10	0.29	0.42	2.65
7 Have you had depressive phases before?	0.56		0.31	0.38	0.49	0.41	0.51
4 Do you currently experience difficulties at work?	0.55		0.30	0.19	0.39	0.26	1.59
2 Do you currently have any family problems and/or difficulties in your romantic relationship?	0.54		0.30	0.29	0.46	0.28	0.91
3 Do you currently have difficulties with friends and acquaintances?	0.51		0.26	0.15	0.36	0.23	1.90
8 Are you taking medication to treat any mental illnesses (psychopharmacological drugs)?	0.46		0.21	0.09	0.28	0.21	2.90
1 Do you suffer from frequently occurring pain?	0.39		0.15	0.36	0.48	0.13	0.59
6 Are you burdened by raising children?	0.35		0.12	0.10	0.29	0.22	2.73
DESY-PAT-2	Factors		h^2	M	SD	r_{it}	V
Items	1 (depressive cognition)	2 (suicidality)	3 (symptoms of fatigue)				
4 In the last 2 weeks, have you had more problems concentrating than usual?	0.80	−0.07	0.14	0.74	0.35	0.48	0.64
5 In the last 2 weeks, have you been ruminating more than usual?	0.78	−0.05	0.14	0.73	0.36	0.48	0.57
17 In the last 2 weeks, have you been more irritable than usual?	0.72	0.00	0.07	0.60	0.23	0.42	1.30
7 In the last 2 weeks, have you felt guilty?	0.71	0.17	−0.16	0.52	0.21	0.40	1.45
6 In the last 2 weeks, have you found decision-making more challenging than usual?	0.64	0.06	0.16	0.63	0.17	0.38	1.72
1 In the last 2 weeks, have you felt down and/or sad often?	0.58	0.21	0.23	0.78	0.35	0.48	0.62
2 In the last 2 weeks, have you had significantly less pleasure in things you usually like to do?	0.55	0.39	0.10	0.82	0.24	0.43	1.22
16 In the last 2 weeks, have you felt like you were failing?	0.51	0.42	0.04	0.73	0.23	0.42	1.87
18 In the last 2 weeks, have you been concerned about things or situations that usually do not bother you?	0.51	0.01	0.23	0.48	0.24	0.43	1.22
19 In the last 2 weeks, have you felt like life is not worth living?	−0.21	1.00	0.19	0.96	0.05	0.22	3.99
20 In the last 2 weeks, have you thought you would rather be dead?	0.07	0.90	−0.07	0.83	0.04	0.20	4.65
14 In the last 2 weeks, have you felt like everything is hopeless?	0.25	0.84	−0.07	0.92	0.10	0.31	2.56
15 In the last 2 weeks, have you felt like everything is meaningless?	0.25	0.82	−0.01	0.89	0.09	0.28	2.88
8 In the last 2 weeks, have you felt lonely?	0.10	0.40	0.34	0.50	0.21	0.41	1.42
11 In the last 2 weeks, have you felt tired and/or exhausted more often than usual?	0.07	−0.15	0.95	0.88	0.48	0.50	0.10
12 In the last 2 weeks, have you felt listless and without energy?	0.00	0.16	0.88	0.91	0.34	0.47	0.68
13 In the last 2 weeks, has everything been more stressful for you than usual?	0.12	0.00	0.75	0.69	0.35	0.48	0.62
10 In the last 2 weeks, did you find everyday activities (eg, getting up, eating, going to work) more difficult to perform than usual?	0.08	0.17	0.72	0.74	0.30	0.46	0.88
3 In the last 2 weeks, have you had less interest in your activities than usual?	0.43	0.11	0.47	0.77	0.26	0.44	1.10
9 In the last 2 weeks, have you found yourself reducing your social encounters?	0.22	0.17	0.38	0.43	0.21	0.40	1.45
Highest loadings are printed bold.							
*Factor was tested independently.							
DESY-PAT-1, Questionnaire for the assessment of DEpression SYmptoms in Primary Care, self-rating part for patients 1; DESY-PAT-2, Questionnaire for the assessment of DEpression SYmptoms in Primary Care, self-rating part for patients 2; h^2 , communality score; M, mean; ML, maximum-likelihood; r_{it} , discriminatory power; V, skewness.							

Table 3 ML-factor analysis with loadings of the DESY-GP, descriptive values

DESY-GP	Factor					
Items	1 (depression symptoms)	h^2	M	SD	r_{it}	V
8 Does this patient show signs of joylessness and/or loss of interest?	0.98	0.95	0.15	0.36	0.77	1.97
9 Does this patient show signs of dejection, melancholy and/or hopelessness?	0.96	0.93	0.21	0.41	0.79	1.45
1 Do I have the impression that this patient is depressed?	0.93	0.87	0.22	0.41	0.78	1.37
6 Has this patient shown signs of social withdrawal?	0.91	0.83	0.15	0.36	0.70	1.93
11 Does this patient show signs of impaired concentration?	0.88	0.78	0.18	0.39	0.70	1.63
7 Has this patient shown signs of worrying about the future?	0.88	0.77	0.22	0.42	0.69	1.34
3 Does this patient show signs of reduced resilience in their daily life?	0.86	0.74	0.35	0.48	0.63	0.61
	2(medical history/external factors)					
10 Does this patient show signs of sleep disorders?	0.85	0.73	0.21	0.41	0.47	1.45
5 Has this patient mentioned family problems?	0.80	0.63	0.23	0.42	0.47	1.26
4 Has this patient mentioned work-related problems?	0.56	0.31	0.14	0.35	0.27	2.01
2 Do I agree that the patient's reason for the encounter sufficiently explains the symptoms presented? (inverted)	0.55	0.29	0.11	0.31	0.30	2.54
15 Do I notice anything else unusual regarding depression?	0.52	0.27	0.12	0.32	0.28	2.30
13 Does this patient have any close relatives with mental illness?	0.45	0.20	0.13	0.34	0.24	2.20
14 Does this patient have any relevant physical illnesses?	0.30	0.09	0.43	0.49	0.19	0.28
Universal item: 12 Does this patient have a history of depressive phases?	–	–	0.35	0.48	–	1.97
Factors were tested independently. DESY-GP, Questionnaire for the assessment of DEpression SYmptoms in Primary Care, external rating part for general practitioners; h^2 , communality score; M, mean; r_{it} , discriminatory power; V, skewness.						

the factor ‘medical history/external factors’, for which we found a relatively low internal consistency (Cronbach’s $\alpha=0.59$). One possible explanation for the low consistency is the rather broad range of external risk factors for depression,⁵³ which may be difficult to capture in a single consistent factor. Nevertheless, the factor ‘medical history/external factors’ remains important for the DESY-GP as it reflects GP-specific heuristics.^{35 38}

Furthermore, our findings implicate a high convergent validity of the DESY-PC, as its correlation with the validated depression questionnaire PHQ-9 is significant. However, the DESY-GP is less associated with the PHQ-9 than the DESY-PAT ($r=0.45$ compared with $r=0.81$). This indicates as well that the DESY-GP possibly measures a different

aspect of depression, which is essential for the general practice context. The DESY-PAT, on the other hand, correlates highly with the PHQ-9 ($r=0.81$), reflecting the similarity of the content of the two questionnaires. The DESY-PAT-1 shows a lower correlation with the PHQ-9 than the DESY-PAT-2 ($r=0.57$ compared with $r=0.81$). This difference in correlation with the PHQ-9 reflects the fact that the DESY-PAT-1 captures environmental and contextual factors for depression that are not captured by the PHQ-9, but which can be a useful addition for effective diagnostic decision-making in general practice. There are already many validated depression questionnaires, such as the PHQ-9 or the Hospital Anxiety and Depression Scale.⁴¹ Therefore, a detailed investigation of

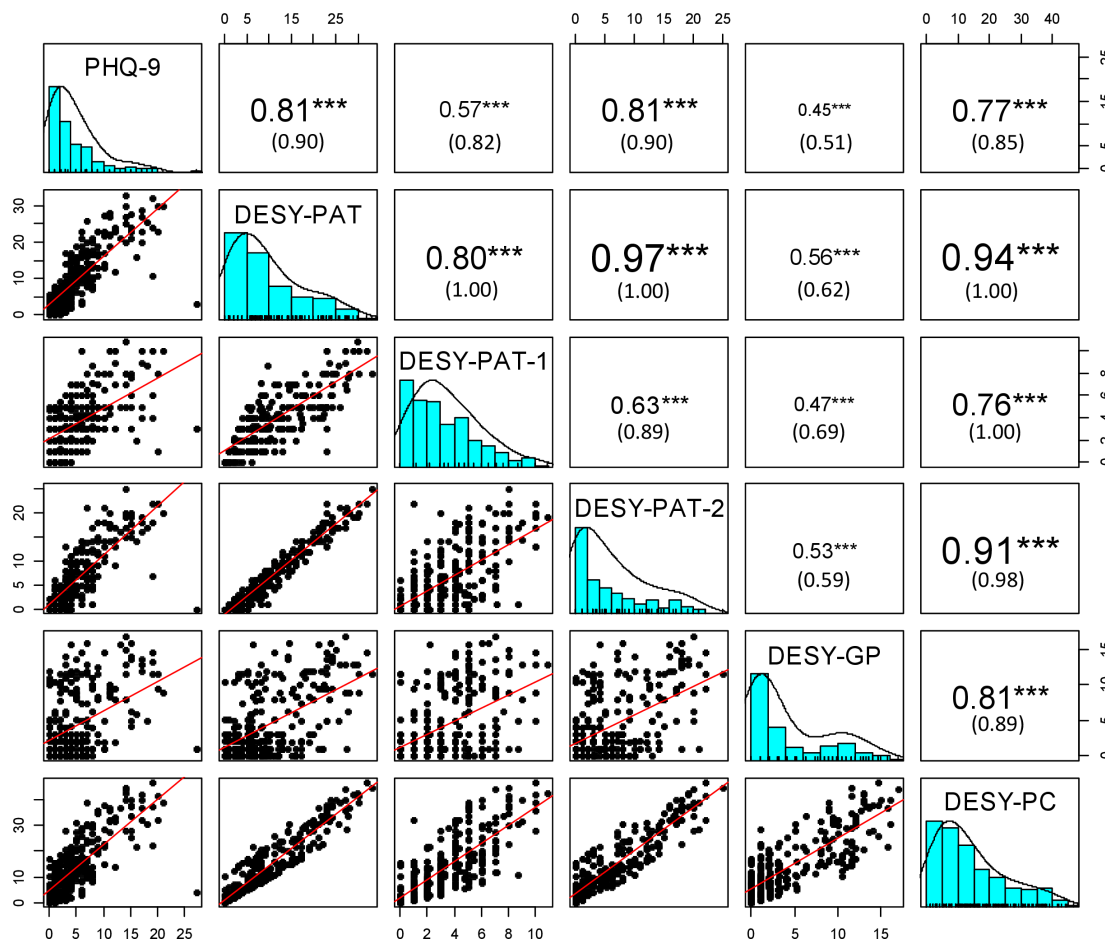


Figure 2 Correlations of DESY-PC and PHQ-9. The values in brackets are the values corrected for attenuation. The numbers were set to one if they exceeded this value. *** $p < 0.001$. DESY-GP, Questionnaire for the assessment of DEpression SYmptoms in Primary Care, external rating part for general practitioners; DESY-PAT, Questionnaire for the assessment of DEpression SYmptoms in Primary Care, self-rating part for patients; DESY-PC, questionnaire with combined questionnaire parts for general practioners and patients; PHQ-9, Patient Health Questionnaire 9.

the diagnostic accuracy of the DESY-PC and all its parts should be carried out using standardised clinical interviews as a reference standard to justify its use as a new symptom-based questionnaire that is adapted to the primary care setting and takes into account the patient's perspectives. If no additional diagnostic use of all parts can be demonstrated, the DESY-PAT-1 and the DESY-GP could be used in addition to already established depression questionnaires to collect contextual information. The high correlation of the DESY-PAT-2 with the PHQ-9 could be an indication of similarity between the two questionnaires and thus partially deprive the DESY-PC of its justification. However, a follow-up study investigates whether the new questionnaire improves the accuracy of diagnostic decision-making in primary care and captures additional information (German Clinical Trials Registry ID: DRKS00031581). A positive finding could be an indicator of the superiority of the DESY-PAT-2 over other validated symptom-based depression questionnaires.

As the DESY-PC is adapted to the primary care setting, it could be used as an improved diagnostic aid for general practice patients who are considered to be at increased risk of depression. It could represent an interesting

alternative to the screening approach of common depression questionnaires.

Strengths and limitations

A strength of the study is that the questionnaire was developed with the help of numerous experts from general practices, psychiatric clinics and patients so that a broad view of the illness of depression is represented. As a resulting innovation, the new DESY-PC questionnaire includes both external and self-report measures. Previous studies have shown that self-assessment is subject to bias and that the inclusion of a clinician's assessment can improve the accuracy of the diagnosis.⁵⁶ In this light, the diagnostic and classification system embedded in World Organization of Family Doctors International Classification for Primary Care (ICPC-3) follows a very similar approach which emerges from the experience of primary care consultations and explicitly includes both GP and patient perspectives.⁵⁷ In contrast to previous editions (ICPC-1 and ICPC-2), there is a shift from a strictly medical or disease-based approach to care to a more person-centred approach. The new questionnaire similarly covers the perspectives of both GPs

and patients. This approach is in line with the ICPC-3 recommendation that better diagnostic decision-making in primary care is achieved by including both perspectives.⁵⁷

Additionally, the closed forced response format (yes/no) of the DESY-PC represents an advantage as it could avoid problems arising from using a middle response category.⁵⁸

However, there are several limitations. In this study, it was not tested whether the DESY-PC identifies depression more accurately than commonly used depression questionnaires. We used the PHQ-9 as the only validated depression screening instrument for comparison. Therefore, in further investigations on the diagnostic accuracy of the new questionnaire, its performance should be compared with an already validated questionnaire regarding one confirmed depression diagnosis. A reference standard like the SCID interview (Structured Clinical Interview for DSM Disorders) should be applied to confirm or rule out a diagnosis. In this way, the sensitivity and specificity of the new two-part questionnaire can be tested and compared with other commonly used depression questionnaires.

A further limitation of our findings might be that we developed our questionnaire with motivated GPs and patients who regularly participate in scientific studies and research projects. These GPs and patients could be more reflective and prone to critical thinking than the average GP and their patients. It remains unclear to what extent this fact influenced the internal consistency of the questionnaire. Additionally, as participation during the validation phase was voluntary, there might have been a selection bias towards more motivated patients. This circumstance may have artificially altered the ratio of depressed to non-depressed patients, as one of these patient groups may be more likely to refuse to participate in the study than the other. Furthermore, patient self-rating questionnaires have the general limitation that patients tend to answer questions influenced by social desirability. However, we accounted for this limitation by implementing an external rating questionnaire for GPs in the DESY-PC.

On a practical level, it remains to be seen how the new questionnaire can be used in primary care and elsewhere. It needs to be clarified whether the questionnaire is to be used only for those suspected of having a depressive disorder or for all primary care patients. Besides, most questionnaires, like the PHQ-9, have a specific cut-off value that indicates a depression diagnosis. For the new questionnaire, no such cut-off exists so far. Future research needs to investigate how a sum score is formed, whether it is weighted and whether all items are equally included in the sum score.

Finally, applying confirmatory and exploratory factor analyses using the same sample is problematic. Thus, the found factor structure must be cross-validated in future studies with a different sample.

CONCLUSION

The new DESY-PC questionnaire combines psychiatric criteria, the patient's perspective and GP heuristics. The questionnaire extends the standard criteria for depressive symptoms and provides additional insight for diagnostic decision-making in general practice. During the development process of the questionnaire, the thought processes and heuristics of GPs, as well as the perspective of their patients, were carefully considered, tailoring the questionnaire for the general practice setting. Factor analysis revealed an easy-to-interpret two-factor (DESY-GP) and four-factor (DESY-PAT) structure of the questionnaire. Overall, the new DESY-PC questionnaire considers both standard diagnostic criteria and diagnostic approaches from general practice, representing an innovative extension of existing diagnostic tools for primary care patients.

Author affiliations

¹Institute of General Practice and Health Services Research, Department of Clinical Medicine, TUM School of Medicine and Health, Technical University of Munich, Munich, Germany

²Department Psychology, LMU Munich, Munich, Germany

³Institute of AI and Informatics in Medicine, TUM School of Medicine and Health, Technical University of Munich, Munich, Germany

⁴Institute of General Practice and Family Medicine, University Hospital of the LMU Munich, Munich, Germany

⁵DFG-Graduiertenkolleg POKAL (DFG-GrK 2621/POKAL-Kolleg), Munich, Germany

Acknowledgements We thank the POKAL-Study-Group (Prädiktoren und Klinische Ergebnisse bei depressiven Erkrankungen in der hausärztlichen Versorgung (POKAL, DFG-GrK 2621)) which consists of the following principal investigators: Tobias Dreischulte, Peter Falkai, Jochen Gensichen, Peter Henningsen, Markus Bühner, Caroline Jung-Sievers, Helmut Krcmar, Karoline Lukaschek, Gabriele Pitschel-Walz and Antonius Schneider. The following doctoral students are as well members of the POKAL-Study-Group: Jochen Vukas, Puya Younesi, Clara Teusen, Feyza Gökce, Victoria von Schrottenberg, Petra Schönweger, Hannah Schillock, Jonas Raub, Philipp Reindl-Spanner, Lisa Pfeiffer, Lukas Kaupe, Carolin Haas, Julia Eder, Vita Brisnik, Constantin Brand, Katharina Biersack, Marie Emilia Vogel and Christopher Ebert. The study was performed for the PhD thesis of Clara Teusen at the Medical Faculty of the Technical University Munich. We want to thank the patient representative from the POKAL-Study-Group advisory board, who advised us on the questionnaire and its application. Furthermore, we would like to thank the GPs and their patients who supported us in this project during the development of the questionnaire and as well during the evaluation phase. Three of the participating practices are members of the Bavarian practice-based research network (BayFoNet, <https://bayfonet.de>).

Contributors AS had the study idea. CT prepared the study protocol, took over data collection, wrote the first draft of the manuscript and was involved in data analysis. MB and AH performed statistical analysis and were involved in manuscript preparation. KL, VvS and JG were involved in reviewing the manuscript. AS was substantially involved in study design and manuscript preparation. CT is the author acting as guarantor and is responsible for the conduct of the study and the decision to publish.

Funding This work was supported by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft, <https://www.dfg.de/>) (grant number GrK 2621).

Disclaimer The funding body had no role in the study design; collection, analysis and interpretation of data; or the preparation, review or approval of the manuscript.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and the Medical Ethics Committee of the Technical University Munich/University Hospital Klinikum rechts

der Isar approved our study (169/21 S-EB, 63/22 S-KK). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. We are committed to making data available to individual interested researchers on reasonable request after approval by our data protection officer and provided we remain within the legal framework.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Clara Teusen <http://orcid.org/0000-0003-2947-355X>

Antonius Schneider <http://orcid.org/0000-0002-2847-8626>

REFERENCES

- McPherson S, Armstrong D. General practitioner management of depression: a systematic review. *Qual Health Res* 2012;22:1150–9.
- Parslow RA, Jorm AF. Who uses mental health services in Australia? An analysis of data from the national survey of mental health and wellbeing. *Aust N Z J Psychiatry* 2000;34:997–1008.
- Gaebel W, Kowitz S, Fritze J, et al. Use of health care services by people with mental illness: secondary data from three statutory health insurers and the German statutory pension insurance scheme. *Dtsch Arztebl Int* 2013;110:799–808.
- Stallman A, Sheeran N, Boschen M. A qualitative exploration of general practitioners' treatment decision-making for depressive symptoms. *Med Decis Making* 2023;43:498–507.
- Trautmann S, Beesdo-Baum K. The treatment of depression in primary care. *Dtsch Arztebl Int* 2017;114:721–8.
- Cape J, McCulloch Y. Patients' reasons for not presenting emotional problems in general practice consultations. *Br J Gen Pract* 1999;49:875–9.
- Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009;374:609–19.
- Spangenberg L, Forkmann T, Brähler E, et al. The association of depression and multimorbidity in the elderly: implications for the assessment of depression. *Psychogeriatr* 2011;11:227–34.
- Stahmeyer JT, Mörtens C, Eidt-Koch D, et al. The state of care for persons with a diagnosis of depression. *Dtsch Arztebl Int* 2022;119:458–65.
- Davidson AS, Fosgerau CF. What is depression? Psychiatrists' and gps' experiences of diagnosis and the diagnostic process. *Int J Qual Stud Health Well-being* 2014;9:24866.
- Gutiérrez-Rojas L, Porras-Segovia A, Dunne H, et al. Prevalence and correlates of major depressive disorder: a systematic review. *Braz J Psychiatry* 1999;42:657–72.
- Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 2011;21:655–79.
- Mack S, Jacobi F, Beesdo-Baum K, et al. Functional disability and quality of life decrements in mental disorders: results from the mental health module of the German health interview and examination survey for adults (Dogs1-MH). *Eur Psychiatry* 2015;30:793–800.
- Murphy JA, Byrne GJ. Prevalence and correlates of the proposed DSM-5 diagnosis of chronic depressive disorder. *J Affect Disord* 2012;139:172–80.
- Steinert C, Hofmann M, Kruse J, et al. The prospective long-term course of adult depression in general practice and the community. A systematic literature review. *J Affect Disord* 2014;152:154:65–75.
- Markkula N, Härkänen T, Nieminen T, et al. Prognosis of depressive disorders in the general population- results from the longitudinal Finnish health 2011 study. *J Affect Disord* 2016;190:687–96.
- van Wijngaarden B, Schene AH, Koeter MWJ. Family caregiving in depression: impact on caregivers' daily life, distress, and help seeking. *J Affect Disord* 2004;81:211–22.
- Caan W. The global crisis of depression: the low of the 21st century? *Perspect Public Heal* 2015;135:62.
- Vos T, Allen C, Arora M. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet* 2016;388:1545–602.
- World Health A. *Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level: report by the secretariat*. Geneva: World Health Organization, 2012.
- Engels A, König H-H, Magaard JL, et al. Depression treatment in Germany - using claims data to compare a collaborative mental health care program to the general practitioner program and usual care in terms of guideline adherence and need-oriented access to psychotherapy. *BMC Psychiatry* 2020;20:591.
- Carey M, Jones K, Meadows G, et al. Accuracy of general practitioner unassisted detection of depression. *Aust N Z J Psychiatry* 2014;48:571–8.
- Zhang R, Peng X, Song X, et al. The prevalence and risk of developing major depression among individuals with subthreshold depression in the general population. *Psychol Med* 2023;53:3611–20.
- Joffres M, Jaramillo A, Dickinson J, et al. Recommendations on screening for depression in adults. *CMAJ* 2013;185:775–82.
- National Institute for Health and Care Excellence (NICE). Depression in adults: treatment and management. NICE Guideline; 2022. Available: <https://www.nice.org.uk/guidance/ng222> [Accessed 24 Oct 2023].
- Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale versorgungsleitlinie unipolare depression – leitlinienreport, version 3.0.2022. [german national disease management guideline (NDMG) on unipolar depression – guideline report, version 3.0.2022]. 2022. Available: www.leitlinien.de/depression
- Mitchell AJ, Yadegarfar M, Gill J, et al. Case finding and screening clinical utility of the patient health questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies. *BJPsych Open* 2016;2:127–38.
- Gilbody S, Richards D, Brealey S, et al. Screening for depression in medical settings with the patient health questionnaire (PHQ): a diagnostic meta-analysis. *J Gen Intern Med* 2007;22:1596–602.
- Thombs BD, Coyne JC, Cuijpers P, et al. Rethinking recommendations for screening for depression in primary care. *CMAJ* 2012;184:413–8.
- Costantini L, Pasquarella C, Odone A, et al. Screening for depression in primary care with patient health questionnaire-9 (PHQ-9): a systematic review. *J Affect Disord* 2021;279:473–83.
- Arroll B, Goodyear-Smith F, Crengle S, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Ann Fam Med* 2010;8:348–53.
- Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *Br J Gen Pract* 2007;57:144–51.
- Thombs BD, Markham S, Rice DB, et al. Screening for depression and anxiety in general practice. *BMJ* 2023;382:1615.
- Sinnema H, Terluin B, Volker D, et al. Factors contributing to the recognition of anxiety and depression in general practice. *BMC Fam Pract* 2018;19:99.
- Schumann I, Schneider A, Kanter C, et al. Physicians' attitudes, diagnostic process and barriers regarding depression diagnosis in primary care: a systematic review of qualitative studies. *Fam Pract* 2012;29:255–63.
- van Weel-Baumgarten EM, van den Bosch WJ, van den Hoogen HJ, et al. The validity of the diagnosis of depression in general practice: is using criteria for diagnosis as a routine the answer. *Br J Gen Pract* 2000;50:284–7.
- Wilhelm KA, Finch AW, Davenport TA, et al. What can alert the general practitioner to people whose common mental health problems are unrecognized. *Med J Aust* 2008;188:S114–8.
- Pilars de Pilar M, Abholz H-H, Becker N, et al. How do general practitioners deal with patients they do not consider to be depressed but who are classified as such according to the PHQ-9. *Psychiatr Prax* 2012;39:71–8.
- André M, Borgquist L, Mölsted S. Use of rules of thumb in the consultation in general practice - an act of balance between the individual and the general perspective. *Fam Pract* 2003;20:514–9.
- Teusen C, Hapfelmeier A, von Schrottenberg V, et al. Combining the GP's assessment and the PHQ-9 questionnaire leads to more

- reliable and clinically relevant diagnoses in primary care. *PLOS ONE* 2022;17:e0276534.
- 41 Cameron IM, Cardy A, Crawford JR, *et al.* Measuring depression severity in general practice: discriminatory performance of the PHQ-9, HADS-D, and BDI-II. *Br J Gen Pract* 2011;61:e419–26.
 - 42 Oõpik P, Aluoja A, Kalda R, *et al.* Screening for depression in primary care. *Fam Pract* 2006;23:693–8.
 - 43 Terluin B, Brouwers EPM, van Marwijk HWJ, *et al.* Detecting depressive and anxiety disorders in distressed patients in primary care; comparative diagnostic accuracy of the four-dimensional symptom questionnaire (4DSQ) and the hospital anxiety and depression scale (HADS). *BMC Fam Pract* 2009;10:58.
 - 44 Sharp LK, Lipsky MS. Screening for depression across the lifespan: a review of measures for use in primary care settings. *Am Fam Physician* 2002;66:1001–8.
 - 45 Smarr KL, Keefer AL. Measures of depression and depressive symptoms: beck depression inventory-II (BDI-II), center for epidemiologic studies depression scale (CES-D) geriatric depression scale (GDS), hospital anxiety and depression scale (HADS), and patient health questionnaire-9. *Arthrit Care Res* 2011;63:454–66.
 - 46 Willis GB, Royston P, Bercini D. The use of verbal report methods in the development and testing of survey questionnaires. *Appl Cogn Psychol* 1991;5:251–67.
 - 47 Kroenke K, Spitzer RL, Williams JBW, *et al.* The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *Gen Hosp Psychiatry* 2010;32:345–59.
 - 48 Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the patient health questionnaire (PHQ-9): a meta-analysis. *CMAJ* 2012;184:191–6.
 - 49 William R. Psych: procedures for psychological, psychometric, and personality research. R package version 2.3.12. Evanston, Illinois Northwestern University; 2023. Available: <https://CRAN.R-project.org/package=psych>
 - 50 Fabrigar LR, Wegener DT, MacCallum RC, *et al.* Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods* 1999;4:272:272–99.
 - 51 lavaan RY. An R package for structural equation modeling. *J Stat Softw* 2012;48:1–36.
 - 52 World Health Organization (WHO). International classification of diseases, eleventh revision (ICD-11). Licensed under creative commons attribution-NoDerivatives 3.0 IGO licence (CC BY-ND 3.0 IGO); 2019. Available: <https://icd.who.int/browse11>
 - 53 Herrman H, Patel V, Kieling C, *et al.* Time for united action on depression: a lancet–world psychiatric association commission. *Lancet* 2022;399:957–1022.
 - 54 Wittchen HU, Lieb R, Wunderlich U, *et al.* Comorbidity in primary care: presentation and consequences. *J Clin Psychiatry* 1999;60 Suppl 7:29–36.
 - 55 Baik S-Y, Bowers BJ, Oakley LD, *et al.* The recognition of depression: the primary care clinician's perspective. *Ann Fam Med* 2005;3:31–7.
 - 56 Cuijpers P, Li J, Hofmann SG, *et al.* Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clin Psychol Rev* 2010;30:768–78.
 - 57 Napel HT, van Boven K, Olagundoye OA, *et al.* Improving primary health care data with ICPC-3: from a medical to a person-centered perspective. *Ann Fam Med* 2022;20:358–61.
 - 58 Bortz J, Döring N. Forschungsmethoden und evaluation für human- und sozialwissenschaftler: limitierte sonderausgabe. [Research methods and evaluation for human and social scientists: limited special edition.]. Springer-Verlag, 2006. Available: <http://link.springer.com/10.1007/978-3-540-33306-7>