PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Development and psychometric evaluation of a questionnaire for	
	the assessment of depression in primary care: A cross-sectional	
	study	
AUTHORS	Teusen, Clara; Bühner, Markus; Hapfelmeier, Alexander; von	
	Schrottenberg, Victoria; Linde, Klaus; Gensichen, Jochen;	
	Schneider, Antonius	

VERSION 1 – REVIEW

REVIEWER	Dowrick, CF
	University of Liverpool
REVIEW RETURNED	25-Jan-2024

GENERAL COMMENTS	This paper has the valuable objective of combining screening for depression with GP heuristics on the context in which a depressive episode may be present. The authors argue, correctly, that existing instruments (such as PHQ9) focus exclusively on symptoms and do not consider context.
	They describe a well-considered process of generating initial questions, modifying these in the light of think-aloud cognitive interviews with GPs, psychiatrists/psychotherapists and - importantly - patients, and then testing the refined questionnaires in routine practice with GPs and patients. They undertake exploratory factor analysis to identify key factors in GP and patient questionnaires, and note high correlation between their 29 item patient symptom checklist and PHQ9. They note relevant methodological and practical limitations of this research.
	I have the following concerns.
	1. There are some linguistic infelicities in the translation of the German text into English, for example 'does this patient make a depressive/irritated impression on me?' might be better rendered as 'do I have the impression that this patient is depressed/irritated?' I would encourage the authors to review these transliterations with another native English speaker fluent in German.
	2. It would be helpful to have further information on how the initial set of questions was chosen and defined. Were they based on a formal classification system, such as ICD-10, and/or on more

extensive psychiatric formulation and/or on the clinical experience of the research team?
3. The paper would benefit from further discussion of the need for a new symptom-based questionnaire. Given the close correlation why PHQ-9, why not use that instrument - or other validated insgtruments such as Beck Depression Inventory or Hospital Anxiety and Depression Scale - and supplement it with the new information provided by the DESY-PAT 2 and the DESY-GP?
4. It would also be useful to see discussion of the relevance of the diagnostic and classification system embedded in WONCA's International Classification for Primary Care (ICPC-3) which emerges from the experience of primary care consultations and explicitly includes both GP and patient perspectives.

REVIEWER	Beckers, Thijs
	Hogeschool Arnhem Nijmegen University of Applied Science,
	Research Group Social Psychiatry and Mental Health Nursing
REVIEW RETURNED	31-Jan-2024

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript.
	Overall this study has sufficient merit and quality for publication, however I do have some remarks that need addressing: - The introduction starts right away with the background of depression. I feel the mansucrupts' readability would improve if you would start with a short section on the reason for conducting this study.
	The background section is mostly thorough and well-written.The methods section is clear and the statistical analysis is
	described in sufficient detail
	 I appreciatie the description of the changes in the instrument in each stage.
	- The limitations of this study are considerable, mainly due to not comparing the instrument to a 'gold standard'. However, the author describe these limitations thoroughly, which I consider more than
	sufficient to acknowledge these limitations.
	 I do not see why the anonymised dataset could not be provided in an open access repository.

REVIEWER	Krug, Katja University Hospital Heidelberg, Dept. of General Practice and
	Health Services Research
REVIEW RETURNED	16-Feb-2024

GENERAL COMMENTS	Thank you for your work and the rigorous approach to develop and validate a diagnostic tool for depression in general practice patients. I enjoyed reading the manuscript and am looking forward to your further refining and validating the tool for practical use.
	Just a few suggestions for the manuscript: 1. In the abstract, I wondered about the difference between the patient versions DESY-PAT-1 and DESY-PAT-2. It is perfectly

clear in the main text but I would suggest adding a short description or at least a keyword to the respective version in the abstract.

2. The questionnaire is supposed to be applied to patients with suspected depressive symptoms. Were only those "target patients" contacted for the cross-sectional study? It sounds like there was no such distinction and any patient in the waiting room was approached (I. 152) or was there a pre-screening by the GPs to approach the intended population?

3. I. 206: How were the scales prepared for Pearson correlation? The items of a factor were probably summed up; please describe it explicitly.

Thank you!

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. CF Dowrick, University of Liverpool

Comments to the Author:

This paper has the valuable objective of combining screening for depression with GP heuristics on the context in which a depressive episode may be present. The authors argue, correctly, that existing instruments (such as PHQ9) focus exclusively on symptoms and do not consider context.

They describe a well-considered process of generating initial questions, modifying these in the light of think-aloud cognitive interviews with GPs, psychiatrists/psychotherapists and - importantly - patients, and then testing the refined questionnaires in routine practice with GPs and patients. They undertake exploratory factor analysis to identify key factors in GP and patient questionnaires, and note high correlation between their 29 item patient symptom checklist and PHQ9. They note relevant methodological and practical limitations of this research.

We sincerely thank you for dedicating your time to reviewing our manuscript and providing us with your thoughtful and constructive comments. We have carefully considered each of your points and have made the necessary revisions to the manuscript. Below, you will find our in-line responses to your comments and questions, accompanied by excerpts of the corresponding manuscript changes.

I have the following concerns.

1. There are some linguistic infelicities in the translation of the German text into English, for example 'does this patient make a depressive/irritated impression on me?' might be better rendered as 'do I have the impression that this patient is depressed/irritated?' I would encourage the authors to review these transliterations with another native English speaker fluent in German.

Thank you very much for your valuable advice. We have taken your suggestion seriously and have had the English translation of the questionnaire revised by a native speaker who is fluent in German from the English Writing Centre at the Technical University of Munich. This process was extremely important to improve the English comprehensibility of the questionnaire. We have also accepted your specific suggestions for the wording of the first two items of the DESY-GP ('Does this patient make a depressive/irritated impression on me?'). Apart from that, we hope that the changes in the wording are now more appropriate and more in line with English standards. The wording has been adapted in both parts of the questionnaire. The changes can be found in the Supplementary Material with tracked changes. Additionally, we provide screenshots of the new version with tracked changes below.

Dear colleague,

We ask-would like to ask you to fill outcomplete this questionnaire for depression diagnostics after the consultation with your patient. The following questions are designed intended to help you assess whether or notif the patient you are examining suffers from depression. Try to evaluate answer the following questions by considering using your impression from the last consultation and also your general knowledge of the patient. If none of the options answer alternative seems correct, choose the one that is most likely to be accurate.

		Yes	No.
1.	Do I have the impression thates this patient make a depressive impression on		
	depressed?	_	
2.	Does Do I have the impression that this patient make anis irritated impression		
on m		_	_
3. and s	Do I agree that the patient's the current reason for the consultation encounter sufficiently explains the		
ana <u>s</u>	symptoms form a coherent		
	picturepresented?		
4.	Does this patient show is there a more substantial pain experience than that		
accor	rding-defined to the by		
	medical —findings (e.g. increased complain <mark>tsing</mark>)?		
5.	Does this patient show signs is there evidence of reduced resilience in their	П	П
daily			
6.	Does this patient show signs Is there evidence of increased fatigue and/or		
	ustion?		
7.	Has this patient claimed an abnormal number Are there any abnormalities in ingof attestations or certificates of incapacity		
ыанн	— for workwork incapacity		
	certificates?		
8.	Has this patient mentioned Is there evidence of work-related problems?		
9.	Has this patient mentioned Is there evidence of family problems?		
10.	Has this patient shown signs is there evidence of social withdrawal?		
11.	Has this patient shown signs is there evidence of worrying about the future?		
12.	Does this patient show signs is there evidence of joylessness and/or loss of		П
intere			
13.	Does this patient show signs Is there evidence of dejection, melancholy and/or		
hope	lessness?		
14.	Does this patient show signs Is there evidence of sleep disorders?		
15.	Does this patient show signs is there evidence of impaired concentration?		
16.	Does this patient hHave there ever been history of depressive phases?		
17.	Does this patient have Are there any close relatives with mental illness?		
18.	Does this patient show signs is there evidence of an addiction problem (C2,		
	ne, cannabis,	_	
	_medication,		
	other drugs, media or gambling addiction)?	_	_
19.	Are there Does this patient have any relevant physical illnesses?		
20.	For women: Is-aDoes this patient use hormonal contraceptive-sbeing utilized?		
21.	Do I noticees anything else unusual regarding depression-seem unusual to me?		

We are interested in factors that are often associated with depression. Please answer each question as well as you can. If none of the options answer alternative seems suitable for to you, choose the one that corresponds the most to your situation.

	Yes	No
Do you have any physical illnesses from which you particularly suffer?		
Do you suffer from frequently occurring pain?		
3. Do you currently have any family strainsproblems?		
4. Do you currently have difficulties with friends and acquaintances?		
5. Do you currently experience difficulties in your <u>romantic</u> relationship?		
Do you currently experience difficulties at work?		
7. Do you currently have any financial difficulties?		
Are you burdened by raising children?		
9. Have you had depressive phases before?		
10. Were there any events in your life that were particularly distressing for you?		
11. Have you been or are you receiving treatment for a mental illness?		
12. Are you taking medication in connectionto treat with any mental illnesses (psychopharmacological drugs)?		
13. Are thereDoes anyone in your immediate family have a mental illnesses in your immediate family?		

In the following, we are interested in how you have been feeling lately. The following questions are about the past 2 weeks. Please answer each question as well as you can. If none of the options answer alternative-seems suitable for to you, choose the one that corresponds most to your situation.

<u>—</u> . ·	Yes	No
In the last 2 weeks, have you felt down and/or sad often?		
2. In the last 2 weeks, have you had significantly less pleasure in things you usually like to do?		
3. In the last 2 weeks, have you had less interest in your activities than usual?		
4. In the last 2 weeks, have you had more problems concentrating than usual?		
5. In the last 2 weeks, have you been ruminating more than usual?		
6. In the last 2 weeks, have you found <u>decision-making decisions-more challenging</u> than usual?		
7. In the last 2 weeks, have you felt guilty?		
8. In the last 2 weeks, have you felt lonely?		
 In the last 2 weeks, have you <u>found yourself</u> reducing <u>ed</u>-your social <u>contactsencounters</u>? 		
10. In the last 2 weeks, did you find everyday activities (e.g. getting up, eating, going to work) more difficult to perform than usual?		
	Yes	No
11. In the last 2 weeks, have you been sleeping worse than usual (e.g., disturbed		
trouble		
— falling asleep, trouble staying asleep and/or sleeping through the night, early morning		
awakenings, and/or		
—increased amount		
of sleep)?		

2. It would be helpful to have further information on how the initial set of questions was chosen and defined. Were they based on a formal classification system, such as ICD-10, and/or on more extensive psychiatric formulation and/or on the clinical experience of the research team?

We agree that it is helpful for the reader to have specific information about the origin of the initial items of the questionnaire. To provide further information on how the initial set of questions was chosen and defined, we have revised and added the following sentences in the methods section:

"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".

3. The paper would benefit from further discussion of the need for a new symptom-based questionnaire. Given the close correlation why PHQ-9, why not use that instrument - or other validated insgtruments such as Beck Depression Inventory or Hospital Anxiety and Depression Scale - and supplement it with the new information provided by the DESY-PAT 2 and the DESY-GP?

We appreciate this valuable suggestion. It is a reasonable point to discuss the need for a new symptom-based questionnaire, especially as many validated questionnaires already exist. However, we want to emphasise that during the development of the new questionnaire, the patient's perspectives were considered, with the hypothesis that this might improve diagnostic accuracy. It is necessary to first test the performance of the new questionnaire against already validated depression screeners before we can decide whether additional diagnostic information is captured and whether this new questionnaire is better adapted to the primary care setting. We have added the following sentences to the discussion to address the arguments raised:

"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 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."

4. It would also be useful to see discussion of the relevance of the diagnostic and classification system embedded in WONCA's International Classification for Primary Care (ICPC-3) which emerges from the experience of primary care consultations and explicitly includes both GP and patient perspectives.

Thank you for raising this interesting point. The approach of the WONCA International Classification for Primary Care (ICPC-3) to include a more patient-centred approach in addition to the GP's expert opinion is similar to the approach of the new questionnaire with a questionnaire part for GPs and a self-rating part for their patients. We have therefore included this aspect in the discussion:

"In this light, the diagnostic and classification system embedded in WONCA's (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]."

Reviewer: 2

Mr. Thijs Beckers, Hogeschool Arnhem Nijmegen University of Applied Science

Comments to the Author:

Thank you for the opportunity to review this manuscript.

Overall this study has sufficient merit and quality for publication, however I do have some remarks that need addressing:

Thank you for taking the time to review our manuscript. We appreciate that you consider our approach to be a significant contribution. Below, we address the remarks and comments made.

- The introduction starts right away with the background of depression. I feel the mansucrupts' readability would improve if you would start with a short section on the reason for conducting this study.

We appreciate your suggestion and agree that the beginning of the background on depression is reported somewhat out of context and unrelated to the reason for conducting this study. Therefore, we have replaced paragraphs 1 and 2 to improve readability and to quickly clarify the context and importance of a study to improve depression diagnostics in primary care. The changes can be found in the Main Document with tracked changes. Additionally, we provide screenshots of the new version with tracked changes below.

66 INTRODUCTION 67 Depression is one of the most prevalent mental disorders [1 3]. Various studies have reported a 68 lifetime prevalence of depressive disorders ranging from 12% to 19% [2, 4-6]. Depression has a major 69 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 [7, 8]. Between 70 2005 and 2015, depression rose from the fourth to the third leading cause of disability [9]. Moreover, 71 72 the World Health Organization (WHO) predicts that depression will be the largest burden of disease 73 worldwide by 2030 [10]. Hence, it is particularly important to improve the diagnosis and care of 74 patients with depression and to optimise treatment processes [11]. It is crucial to identify and treat 75 people with depression in the early stages of their illness to prevent chronicity [12]. Besides, proactive 76 management of subthreshold depression can also protect affected individuals from developing major 77 depression [13]. 78 The general practitioner (GP) is usually the first healthcare provider that patients consult [1-34-16]. In 79 most cases, GPs are also the gatekeepers for further diagnostics and treatment of patients with 80 depression [4, 517, 18]. However, identifying depression in primary care can be challenging when only 81 somatic symptoms are reported, and patients do not explicitly mention their depressed mood [619]. 82 In addition to this challenge, the diagnosis of depression in primary care is further complicated by 83 multimorbidity. Somatic complaints often overlap and mask symptoms of depression, so it can be difficult to distinguish between somatic disorders and depression [20, 217, 8]. In any case, the initial 84 diagnosis is essential for subsequent treatment [518, 922]. Thus, it is crucial that GPs follow a guideline-85 86 oriented diagnostic process and treatment, as the majority of patients with depression are only seen 87 in general practice [922, 1023]. 88 In this context, it is important to note that depression is one of the most prevalent mental disorders 89 [11-13-3]. Various studies have reported a lifetime prevalence of depressive disorders ranging from 90 12% to 19% [12, 14-16]. Depression has a major impact on the lives of those who are affected, on their 91 family members, and on their immediate environment. Therefore, it represents a considerable health 92 problem for our society [17, 18]. Between 2005 and 2015, depression rose from the fourth to the third 93 leading cause of disability [19]. Moreover, the World Health Organization (WHO) predicts that 94 depression will be the largest burden of disease worldwide by 2030 [210]. Hence, it is particularly 95 important to improve the diagnosis and care of patients with depression and to optimise treatment 96 processes [211]. It is crucial to identify and treat people with depression in the early stages of their 97 illness to prevent chronicity [212]. Besides, proactive management of subthreshold depression can 98 also protect affected individuals from developing major depression [213].

- The background section is mostly thorough and well-written.

Thank you for your compliment.

- The methods section is clear and the statistical analysis is described in sufficient detail

Thank you for your praise of the statistical analysis section.

- I appreciatie the description of the changes in the instrument in each stage.

We are pleased that you were able to follow the changes to the new questionnaire at each stage and that the explanations were clear.

- The limitations of this study are considerable, mainly due to not comparing the instrument to a 'gold standard'. However, the author describe these limitations thoroughly, which I consider more than sufficient to acknowledge these limitations.

Thank you for raising this important point. We agree with you that an important way to validate the new questionnaire would be to use a gold standard. As a next step, we have addressed this point in a follow-up study in which we validate the newly developed questionnaire against a standardised diagnostic interview (German Clinical Trials Registry ID: DRKS00031581).

- I do not see why the anonymised dataset could not be provided in an open access repository.

We understand your point. Unfortunately, our data is pseudonymised and carries the risk of reidentification. According to German law (derived from the GDPR), making data available in a public repository, and thus sharing it with third parties, would only be possible if appropriate consent had been obtained. Unfortunately, this is not the case and we do not have approval from our data protection officers and data protection lawyers. Nevertheless, we are committed to making data available to individual interested researchers upon reasonable request after approval by our data protection officer and provided we remain within the legal framework. For future studies, we will examine the conditions under which such authorisation can be granted, e.g. by consenting to this type of data processing.

Reviewer: 3

Dr. Katja Krug, University Hospital Heidelberg

Comments to the Author:

Thank you for your work and the rigorous approach to develop and validate a diagnostic tool for depression in general practice patients. I enjoyed reading the manuscript and am looking forward to your further refining and validating the tool for practical use.

Thank you very much for taking the time to review our manuscript! We are glad that your overall impression is positive. Thank you as well for your constructive comments, questions and suggestions. We believe that your input substantially improves our work. Below, you find responses to each of your points and the corresponding changes made to the manuscript.

Just a few suggestions for the manuscript:

1. In the abstract, I wondered about the difference between the patient versions DESY-PAT-1 and DESY-PAT-2. It is perfectly clear in the main text but I would suggest adding a short description or at least a keyword to the respective version in the abstract.

This comment is very helpful - thank you! Due to the limited number of words in the abstract, we have decided to briefly explain the difference between DESY-PAT-1 and DESY-PAT-2 in brackets. The adapted sentence now reads:

"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)."

2. The questionnaire is supposed to be applied to patients with suspected depressive symptoms. Were only those "target patients" contacted for the cross-sectional study? It sounds like there was no such distinction and any patient in the waiting room was approached (I. 152) or was there a pre-screening by the GPs to approach the intended population?

Thank you for raising this point. It seems the procedure was not clearly explained in the methods section. In this study, all patients were approached in the waiting room, regardless of their reason for the encounter with the GP. As part of the evaluation of DESY-PC, it is to be examined whether the questionnaire (despite its length) should be used as a screening instrument or primarily only in cases with suspected depression. The final decision can only be made after a sufficient evaluation of the questionnaire. It should be examined whether the screening approach is useful in the primary care setting or whether a stepped approach would be more appropriate. Follow-up studies could investigate

which items of the new questionnaire are crucial for diagnostic accuracy and could, therefore, possibly be used as filter questions. To make our approach more transparent, we have modified the following sentences in the methods section:

"All patients were approached consecutively (i.e. without pre-selection) on certain days at regular intervals in the general practitioner'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."

3. I. 206: How were the scales prepared for Pearson correlation? The items of a factor were probably summed up; please describe it explicitly.

Thank you for pointing out this missing information. To describe the preparation of the scales for Pearson correlation, we have added the following sentence in the methods section:

"Items within a factor were 0/1 dummy-coded and summed, and corresponding sum scores were used to calculate Pearson correlation coefficients."

VERSION 2 – REVIEW

REVIEWER	Dowrick, CF
	University of Liverpool
REVIEW RETURNED	07-May-2024
GENERAL COMMENTS	The authors have carefully and comprehensively addressed all the
	concerns raised in my initial review.
REVIEWER	Krug, Katja
	University Hospital Heidelberg, Dept. of General Practice and
	Health Services Research
REVIEW RETURNED	21-May-2024
GENERAL COMMENTS	Thank you for addressing all issues raised. I have no further
	suggestions.

VERSION 2 – AUTHOR RESPONSE