

Emoqol 100, an ultra ultra-brief mood case-finding tool: A diagnostic accuracy study

Bruce Arroll¹, Connor F. Mulcahy², Rachel P. Roskvist¹, Vicki Mount¹

¹Department of General Practice and Primary Health Care, University of Auckland, Auckland, New Zealand, ²University of Aberdeen, Aberdeen, Scotland, UK

ABSTRACT

Background: Case-finding for low mood in primary care can be time-consuming using current depression inventories. **Aim:** To assess the diagnostic accuracy, of a single verbally administered question on the emotional quality of life (Emoqol 100), for low mood in patients with symptoms of distress in an ambulatory care setting. **Design and setting:** Eligible patients were consecutive patients seen by one of the authors over 13 months with possible distress/low mood. The index test was the verbally asked Emoqol 100, which is the patient's emotional quality of life now, with 100 being perfect emotional health and 0 being the worst imaginable. The reference standard is the written version of the PHQ-9 with a cut point of ≥ 10 . **Methods:** A retrospective audit of consecutive consultations in a single primary care clinic. **Results:** One hundred two patients were seen during the study period, of which 76 met the eligibility criteria for this audit, and there were 215 test results. For a cut point of < 50 on the Emoqol 100 and the PHQ-9 ≥ 10 the sensitivity was 47% (95% CI 39-54), and the specificity was 93% (95% CI 86-100). The positive predictive value was 95%, and the negative predictive value was 37%. **Conclusion:** This is the first accuracy estimation of the Emoqol 100. It appears to have a high specificity which means when it is positive (< 50) it is a good estimate of a high PHQ-9, i.e. a mood issue probably exists. The test will be helpful for busy primary care clinicians as it takes less than 15 seconds to verbally administer.

Keywords: Clinical audit, depression, primary health care, sensitivity and specificity

Introduction

Major depression is common in primary care with a reported rate of 18.1% over the previous 12 months.^[1] Depressive symptoms are a frequent reason for presentation, and vary in severity; for every one person who satisfies criteria for major depression, there will be three people who experience sub-syndromal symptoms.^[2] Despite not meeting criteria for a diagnosis of major depressive disorder, these symptoms can nonetheless confer significant distress and impairment of social functioning.^[2] Primary care

clinicians are advised to be prudent when considering labelling patients with a psychiatric diagnosis based on a single visit, when they may be seeing the patient on the worst day of their lives. Consideration over time is preferable to determine the symptoms' trajectory and subsequent fulfilment of any diagnostic criteria.^[3]

Case-finding and screening tools for depression are useful in primary care to assist in eliciting symptoms, and assessing their severity.^[4,5] Such tools are generally self-reported assessments and can be limited by patients exaggerating or minimizing symptoms, or misunderstanding of the questions and all take a few minutes or more to complete. Most tools relate to periods of days or weeks, while one evaluates the previous year with a single question on depression that confers a likelihood ratio (LR) positive of 2.3 and a (likelihood ratio negative) LR-ve of 0.16.^[6] While brief, this is not very helpful in terms of post-test probabilities.

Address for correspondence: Dr. Bruce Arroll,
Department of General Practice and Primary Health Care,
University of Auckland, Private Bag 92019, Auckland,
New Zealand 1142.
E-mail: bruce.arroll@auckland.ac.nz

Received: 06-01-2021

Revised: 23-03-2021

Accepted: 01-06-2021

Published: 05-11-2021

Access this article online

Quick Response Code:



Website:
www.jfmpc.com

DOI:
10.4103/jfmpc.jfmpc_40_21

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Arroll B, Mulcahy CF, Roskvist RP, Mount V. Emoqol 100, an ultra ultra-brief mood case-finding tool: A diagnostic accuracy study. J Family Med Prim Care 2021;10:3590-4.

The Patient Health Questionnaire 9 (PHQ-9) has become an increasingly used tool for clinicians in determining the likelihood and degree of depressive symptoms. It contains nine self-assessment questions relating to the patient's perceptions of events over the previous two weeks and can be administered independently or by the clinician. Scores range from 0 to 27, with higher scores indicating more severe symptoms. In practice, the PHQ-9 score may be used to assist clinicians in determining whether patients have mild (score 10-14), moderate (score 15-19) or severe (score >20) symptoms of depression.

Kroenke *et al.*^[7] have developed an ultra-brief tool for depression and anxiety consisting of 4 questions from the PHQ-9 and the GAD-7. However, it still requires clinicians to ask four questions, and that takes time. There is also the PHQ-2 which has a sensitivity and specificity of 86% and 78% for a cut point of greater than 2 and 61% and 92% for a cut point of 3.^[8] Another study looked at the 2 Whooley questions and asked a help question. The two questions have a sensitivity of 96% and specificity of 78%, but the specificity for an additional help question for "help wanted today" was 94%.^[9]

The need for a simple, quick, reliable test to evaluate the severity of depressive symptoms in primary care was identified in a single practice by one of the investigators, who, as well as his patients with mental health issues, began seeing patients by referral from his clinic colleagues and from other clinics for mental health issues. It was apparent that there was a need for a simple way of scoring patients to assess if they were improving or not. The 0 to 100 scale of the Euroqol 5D was considered, but this proved unhelpful as it failed to differentiate between the physical and emotional quality of life.^[10] For example, patients with chronic pain would score around 40 on the Euroqol 5D 100 score but say emotionally they would score 90. A single emotional quality of life question (Emoqol 100) was developed to allow more accurate self-assessment of mood at that particular time. The Emoqol 100 question is: "How is your emotional quality of life now, with 100 being perfect and zero being worst imaginable?" The answer is scored from 0 to 100, with 100 being perfect emotional health, and 0 meaning the worst possible. The Emoqol 100 question is verbal, takes less than 15 seconds to apply, and appears to be generally well understood by patients.

The Emoqol 100 question was used in practice alongside the PHQ-9 score, which was required at each visit for funding purposes, allowing comparison of the two scores. Over time, it became apparent that Emoqol 100 scores under 50 were associated with higher PHQ 9 scores. The investigators conducted a validation audit to determine the sensitivity and specificity of the Emoqol 100 against the PHQ 9. The audit aimed to assess the sensitivity and specificity (likelihood ratios and predictive values) of the Emoqol compared with the PHQ-9 as recorded by patients. This paper has been written according to the STARD statement for diagnostic tests.^[11] Audits (if anonymous and retrospective) do not require Ethics approval by National

Health Ethics Committee. Being a retrospective paper, it was not registered as a diagnostic test study.

Methods

A retrospective audit was conducted over 13 months, at a General Practice clinic. Participants were consecutive patients seen by one of the authors, in whom emotional distress was a key issue. Some of these patients were regular patients, and others were referred to him for a FACT (Focussed Acceptance and Commitment Therapy) Consultation.^[12] Patients were eligible for the audit if they had a recorded PHQ-9 score and an Emoqol 100 score at the same visit. The Emoqol 100 was the index test, and the PHQ-9 was the reference standard. Both tests were done in the same visit, usually around five minutes into the consultation time once the initial complaint had been ascertained. The order of doing the PHQ-9 and the Emoqol 100 was variable. No assistance was offered when patients occasionally had difficulty with the questions of either tool; when this happened, the clinician asked the patient to make the response decision themselves to avoid biasing the results.

It was postulated that an Emoqol 100 score <50 would correspond to a PHQ-9 score indicating significant symptoms of low mood. A recent meta-analysis reported that a PHQ 9 score of ≥ 10 maximized combined sensitivity and specificity overall.^[13] For the purposes of this audit, Emoqol scores of <45, <50, <51 and <60 were considered.

The intended use of the Emoqol 100 is as a quick and straightforward tool that gives a meaningful measure of the patient's distress. This audit was exploratory; first seeking to confirm the clinical impression of a high specificity and second to determine which cut-point on the Emoqol 0 to 100 scale would represent the most useful value for clinicians to assess significant distress. While the clinician had other information such as medication and past history, this did not alter the conduct of either the PHQ-9 or the Emoqol 100 test. The reference test (PHQ-9) was not given blindly to the patients. The sensitivity/specificity and other measures of diagnostic accuracy were calculated using the Catmaker calculator on the Centre for Evidence-Based Medicine at the University of Toronto.^[14] The sample size was determined by the availability of an elective student to assist with data collection and the analysis.

Results

We found 76 patients who had at least one PHQ-9 and one Emoqol 100 recorded at the same visit [Table 1]. One hundred two patients were potentially eligible, but for various reasons, did not have both questionnaires administered (75% inclusion rate) [Figure 1]. The range of ethnicities reflects the general population of the clinic study site.

The median value for the PHQ-9 was 14, which is at the high end of the mild depression range, meaning that a

significant proportion of patients were in the moderate to severe range when assessed. 22 of the 76 patients were on antidepressants.

Table 2 shows the measures of diagnostic accuracy. The pre-test probability for a score of <50 on the Emoqol reflecting a positive PHQ-9 ≥10 (prevalence) for the Emoqol cut points of <50, was 94%. Based on the PHQ-9 ≥10, the highest positive predictive value (those who have a positive Emoqol 100, i.e. low score was 95% for a cut-point of <50 [Table 3]. What this means to a clinician for a patient who scores <50 is 95% likely to have

a PHQ-9 score of ≥10, i.e. have a problem with their mood at that visit. For patients with an Emoqol score of ≥80, only 8/17 patients had a PHQ-9 score of 10 or more. For patients with an Emoqol score of <40, the likelihood ratio is 8.7, which would generate significant changes in post-test probabilities, especially in low prevalence settings [Table 4].

Discussion

Summary

Our results show that for a cut point on the Emoqol 100 of <50, the positive predictive value (PPV) of a positive test (i.e. one below 50) is 95%. This is a high PPV and due to both the high prevalence of mood problems in this sample and the high specificity. If this were applied to an average primary care prevalence of depression, estimated at 5.2%, the positive predictive value would be 32% for an Emoqol <50 and (assuming the sensitivity and specificity are the same in the two settings) which is considered high for case finding in primary care.^[8] This is an example of a Spin result where a high specificity (Sp) rules in a condition when the result is positive. (A Snout is a test with a high sensitivity (Sn) and in that situation a negative finding is a good rule out). The cut point for many tests can be adjusted to extenuate the sensitivity or the specificity, especially where both are not high at a range of cut points.

The Emoqol 100 is the briefest of all mental health tools that we are aware of, and hence we have called it an ultra ultra-brief inventory tool. It is also the only one to measure how the patient is feeling today and has the advantage for patients of not having to read and interpret a paper or verbally asked questionnaire.

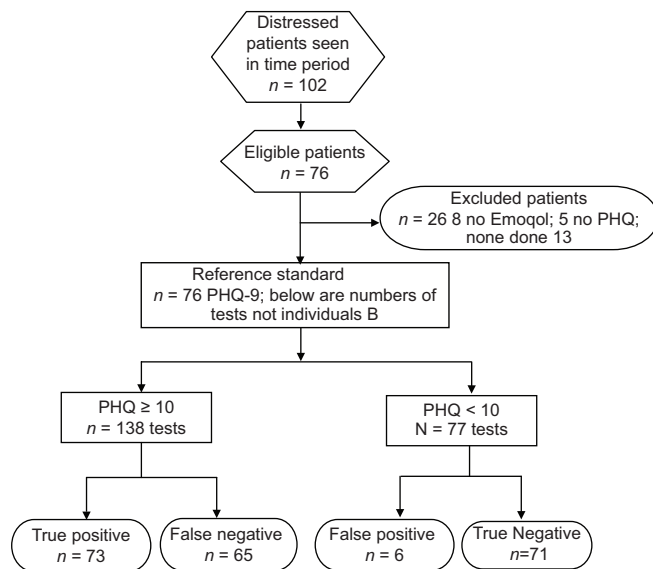


Figure 1: STARD flow diagram for a study of 76 patients with both a PHQ-9 ≤ 10 and Emoqol score (215 records)

Table 1: Baseline measures n=76 and n=215 PHQ 9 measures

| | |
|--------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Age - mean (sd) | 37 years (14.5) |
| Gender | Female: 49 (64%) Male: 27 (36%) |
| Ethnicity | European: 31 (41%) Maori: 17 (22%) Samoan: 7 (9%) Other EU: 6 (8%) Other: 15 (20%) Other: Breakdown (South East Asian: 2 Chinese: 2 Cook Island Maori: 2 EU not defined: 2 Indian: 2 EU: 1 Fijian: 1 Iraqi: 1 Niuean: 1 Other Asian: 1 |
| PHQ range and median | Range: (27-2) Median: 14 |
| Emoqol score range and median | Range: 100 (100-0) Median: 56 |
| # (%) on antidepressants at baseline | 22 (29%) |

Strengths and weaknesses

The strengths and weaknesses of this paper are similar. This is a “real-life” finding and as such is a pragmatic study applied to consecutive patients with distress. It has been validated in a setting where it is intended to be used. The prevalence of low mood will be higher than a consecutive series of patients seen in a usual general practice setting, as this sample included patients who were referred for extra mental health care to one of the authors. It only applies to one practitioner, and as it was a real-life situation, it was not possible to blind the measurement of the reference standard PHQ-9. The gold standard is an inventory and not a gold standard interview. However, the PHQ-9 is becoming the standard tool for primary care assessments, and a briefer option would be helpful. While the PHQ-9 was designed to give a measure of depression it is not a gold standard and in this paper is used as a measure of distress rather than depression.

Comparison with existing literature

A review of 16 case-finding tools for depression reported a range of questions from one question to 30 questions.^[5] The one question on being depressed had a sensitivity of 85% and a specificity of 66%. In contrast, the 20 question CES-D had a

Table 2: Range of cutpoints for Emoqol against a PHQ 9 ≥ 10

| | Cutpoint | Sensitivity | Specificity | Likelihood+ve | Likelihood -ve | PPV |
|--------|---------------|-------------|-------------|---------------|----------------|-----|
| PHQ≥10 | Emoqol 100<45 | 43% | 93% | 5.9 | 0.6 | 93% |
| PHQ≥10 | Emoqol 100<50 | 47% | 93% | 6.3 | 0.6 | 95% |
| PHQ≥10 | Emoqol 100<51 | 56% | 83% | 3.4 | 0.5 | 91% |
| PHQ≥10 | Emoqol 100<60 | 60% | 81% | 3.3 | 0.5 | 91% |

Likelihood ratio+ve=sensitivity/(1-specificity) Likelihood ratio -ve=1-sensitivity/(specificity)

Table 3: Maximum specificity for cut point of PHQ≥10 and Emoqol 100<50% (prevalence=64%)

| Sensitivity | Specificity | Likelihood ratio+ve | Likelihood Ratio -ve | Positive Predictive Value | Negative Predictive Value |
|-------------|--------------|---------------------|----------------------|---------------------------|---------------------------|
| 47% (39-54) | 93% (86-100) | 6.3 (3.1-14.9) | 0.58 (0.49-0.68) | 95% (90-100) | 37% (29-45) |

*(brackets have 95% confidence interval)

Table 4: Likelihood ratios at 3 different cut-points of Emoqol<40, 40 to 60, 61-100

| Emoqol range | Positive PHQ≥10 | PHQ<10 | Likelihood ratio 95% CI |
|--------------|-----------------|--------|-------------------------|
| 0 to 39 | 52 | 2 | 8.7 |
| 40 to 60 | 58 | 17 | 1.2 |
| 61 to 100 | 51 | 35 | 0.5 |

sensitivity of 88% and a specificity of 75%.^[6] The US preventive services task force (2016) recommends screening for depression: they say “Although direct evidence of the isolated health benefit of depression screening in primary care is weak, the totality of the evidence supports the benefits of screening in pregnant and postpartum, and general adult populations, particularly in the presence of additional treatment supports such as treatment protocols, care management, and availability of specially trained depression care providers.” The nine recommended inventories range from 9 to 30 items and take anywhere from 5 to 15 minutes to complete. This is impractical in a standard primary care setting, where consultations may be as short as seven to 10 minutes. The Emoqol is a case-finding tool, not a screening tool and should be used in situations where the pre-test probability of a low mood is higher than usual, i.e. patients complaining of mood symptoms such as sleep difficulties and fatigue.

Implications for research and or practice

We suggest that the Emoqol 100 tool will be of value to primary care clinicians and other clinicians when there is not enough time to do a longer inventory. It takes about 15 seconds and patients do not need to read anything and do not need their glasses. A common situation is where emotional issues are raised near the end of a consultation. A score below 50 would indicate the need for further help and an appropriate intervention made at that visit or a subsequent appointment/referral for additional interventions. Those scoring <40 are highly likely to have a mood issue. For those above 50, the Whooley two questions could be asked; negative answers on the PHQ-2 would rule out depression.^[9] The Emoqol 100 may also be useful for monitoring mood over time.

Further research would necessitate using the Emoqol in different settings and perhaps different mood disorder settings to get a

range of prevalences and seeing what quality of life measures could be assessed. It would also be essential to reproduce the findings in this audit.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

How this fits in is correct

1. This is the first measure of current emotional mood in a single question and can be done in under 15 seconds and can be verbally asked, avoiding the need for reading or the use of glasses.
2. It has a high specificity which means when the score is low, the patient will almost certainly have a mood issue in case finding situations.
3. It will be useful in clinical situations where distress is suspected and there are time limitations such as primary care or secondary care ward rounds.

References

1. Magpie Research Group. The nature and prevalence of psychological problems in New Zealand primary healthcare: A report on mental health and general practice investigation (MaGPIe). *N Z Med J* 2003;116:1171-85.
2. Rapaport MH, Judd LL, Schettler PJ, Yonkers KA, Thase ME, Kupfer DJ, et al. A descriptive analysis of minor depression. *Am J Psychiatry* 2002;159:637-43.
3. Frances A. Allen Frances at preventing over-diagnosis - what opened your eyes. *BMJ* podcast august 31 2018. 2018. Available from: <https://itunes.apple.com/nz/podcast/the-bmj-podcast/id283916558?mt=2&i=1000418438338>.
4. Williams JW Jr, Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed. *JAMA* 2002;287:1160-70.
5. Williams JW Jr, Pignone M, Ramirez G, Perez Stellato C. Identifying depression in primary care: A literature synthesis of case-finding instruments. *Gen Hosp Psychiatry* 2002;24:225-37.
6. Williams JW Jr, Mulrow CD, Kroenke K, Dhanda R,

- Badgett RG, Omori D, *et al.* Case-finding for depression in primary care: A randomized controlled trial. *Am J Med* 1999;106:36-43.
7. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics* 2009;50:613-21.
 8. Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, *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.
 9. Arroll B, Goodyear-Smith F, Kerse N, Fishman T, Gunn J. Effect of the addition of a “help” question to two screening questions on specificity for diagnosis of depression in general practice: Diagnostic validity study. *BMJ* 2005;331:884.
 10. Brooks R, Boye KS, Slaap B. EQ-5D: A plea for accurate nomenclature. *J Patient Rep Outcomes* 2020;4:52.
 11. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ* 2003;326:41-4.
 12. Arroll B. Focussed acceptance and commitment therapy (www.brucearroll.com) 2018.
 13. Levis B, Benedetti A, Thombs BD; DEPRESSION Screening Data (DEPRESSD) Collaboration. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: Individual participant data meta-analysis. *BMJ* 2019;365:11476.
 14. Evidence based medicine toolbox. 2018. Available from: <https://ebm-tools.knowledgetranslation.net/calculator/diagnostic/>.