Emoqol-100: Development and validation of a single question for low mood in primary care

Dahle, Nina Edel; Matthew, Carolyn; Roskvist, Rachel Petronella; Moir, Fiona; Arroll, Bruce

DOI: https://doi.org/10.3399/BJGPO.2023.0011

To access the most recent version of this article, please click the DOI URL in the line above.

Received 18 January 2023
Revised 21 March 2023
Accepted 25 April 2023

© 2023 The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License (http://creativecommons.org/licenses/by/4.0/). Published by BJGP Open. For editorial process and policies, see: https://bjgpopen.org/authors/bjgp-open-editorial-process-and-policies

When citing this article please include the DOI provided above.
Emoqol-100: Development and validation of a single question for low mood in primary care

Authors
Nina Dahle¹ ², Carolyn Mathew³, Rachel Roskvist³, Fiona Moir³, Bruce Arroll³

¹. Centre for Clinical Research, Uppsala University, Falun, Sweden.
². Primary Health Care Center Britsarvet-Grycksbo, County of Dalarna, Falun, Sweden.
³. Department of General Practice and Primary Health Care, University of Auckland, Auckland, New Zealand.

Corresponding author:
Bruce Arroll
ORCID ID: https://orcid.org/0000-0002-0451-9462
Email: bruce.arroll@auckland.ac.nz
Phone: +64-9-9236978

Nina Dahle, MD, Family Medicine resident
Carolyn Mathew, Medical student
Rachel Roskvist MBChB, Senior Lecturer
Fiona Moir, MBBS, PhD Senior Lecturer
Bruce Arroll, MBChB, PhD, Professor

Word count: 2746 words
Abstract

Background
Using current inventories, case-finding for low mood in primary care is important but time-consuming.

Aim
We aimed to validate the diagnostic accuracy of a single question on the emotional quality of life (Emoqol-100) as a measure of depression in symptomatic patients.

Design and setting
This retrospective clinical audit validating the Emoqol-100 compared with the 9-item Patient Health Questionnaire (PHQ-9) and Burns Depression Scale Today (BDST).

Method
Consecutive patients with suspected low mood, seen over 22 months in a single primary care clinic by one of the authors, were eligible for this retrospective audit (n=160). The index test was the verbally asked Emoqol-100; "How is your emotional quality of life now, with 100 being perfect and 0 being the worst imaginable?" The reference standard was the PHQ-9 (n=426 visits) with a cut point of ≥10 or BDST (n=513 visits) with a cut point of ≥6.

Results
The Emoqol-100 range 0-20 had a likelihood ratio (LR) of 25.2 for low mood compared to the BDST and for Emoqol-100 scores of 21-40, 41-60, 61-80 and 81-100 the likelihood ratios were 3.6, 1.7, 0.35 and 0.09 respectively. For the PHQ-9 these were 10.1, 2.9, 1.3, 0.40 and 0.2, respectively. Any score ≤ 60 is associated with a low mood.

Conclusion
The Emoqol-100 appears to have high validity, so when it is low (≤ 60), it is suggestive of a high PHQ-9 or BDST score, and a mood issue probably exists. Emoqol-100 could be helpful for busy primary care and other clinicians.

Keywords: depressive disorder, primary health care, audit, diagnostic tests,
**How this fits in**
This is the first validation study of a single question about the emotional quality of life against the PHQ-9 (The Emoqol-100) for case-finding depression in primary care. It is the first derivation study of the Emoqol-100 compared with the BDST. The Emoqol-100 has a high LR when the score is low e.g. $\leq 60$, meaning the patient is likely to have a mood issue. There are significant clinical changes with the high and low likelihood ratios. For example, with the BDST the pre-test probability is 67% and for a LR of $\leq 20$ the post-test probability of a low mood is 98%. For a LR $\geq 80$ the post-test probability is 15%. These are clinically useful changes in the probability of low mood. It has the potential to be useful in clinical situations where low mood is suspected and time is limited, and a rapid assessment of mood would be clinically useful.

**Introduction**
Depression is usually managed in the primary care setting (1, 2), and the ability of primary care clinicians to effectively diagnose and manage depression is critically important.(3) A meta-analysis indicates that general practitioners (GPs) fail to diagnose over 50% of the depressed patients in their clinic, even if the diagnostic accuracy improves when the GPs meet their patients over an extended period.(4) In contrast, those presenting with somatic symptoms for which no apparent cause can be found are less likely to be recognised.(4) A broad case-finding approach using a short mood inventory test can help GPs correctly identify depressed patients promptly in primary care.(5)

Depression inventories are useful in primary care to assist clinicians to determine the likelihood and degree of depressive symptoms.(6, 7) The Patient Health Questionnaire (PHQ-9) has been identified as one of the most valid.(8, 9) It relates to the patient's symptoms over the previous two weeks.

The Burns Brief Mood Survey is another tool that includes three five-item subscales for assessment of depression, anxiety, and anger during the previous 1-week period. The Burns Brief Mood Survey has been shown to be reliable, with excellent internal consistency.(10) There is also another shorter Burns questionnaire with five questions about the mood today, which we have labelled Burns Depression Scale Today (BDST) to make clear the distinction from the one-week Burns Brief Mood Survey.
The BDST (Supplementary Table) can be used to assess low mood today, and to our knowledge, it has not been validated apart from one conference abstract.\(^{(11)}\)

The need for a quicker and reliable test to evaluate the severity of symptoms of low mood in primary care was identified by one of the investigators. The Emoqol-100 is a single question with a derivation validation conducted against the PHQ-9 \(^{(12)}\); the Emoqol-100 question follows: "How is your emotional quality of life now, with 100 being perfect and zero being the worst imaginable?" The answer is scored verbally from 0 to 100, with 100 being perfect emotional health and 0 meaning the worst imaginable. The Emoqol-100 question is verbally administered, takes less than 15 seconds to apply and interpret and appears to be well understood by most patients.\(^{(12)}\)

The Emoqol-100 was first used in the clinic alongside the PHQ-9. During the Covid-19 pandemic, the BDST was initiated for phone consultations, in addition to the Emoqol-100, as it was easier to administer than the PHQ-9. When time was short, only the Burns inventory was used due to the time needed for the PHQ-9. The PHQ-9 was also administered where possible as it was required for payment purposes, allowing comparison of the three scores.

The first study to investigate the value of a test is called the derivation study, while the second validation study should validate the derivation findings in a different population. The studies should be performed according to the STARD statement \(^{(13)}\) and the final test of a diagnostic test or model should investigate whether it is accurate and generalisable enough for the purpose for which it was derived.\(^{(14)}\) Criterion validity involves comparison with a gold standard, which is called concurrent validity when the comparison is made simultaneously.\(^{(15)}\) These validation/derivation studies are ideally done in the settings where the diagnostic test will be used.\(^{(16)}\)

This paper aims to validate the findings in the derivation study of Emoqol-100 and PHQ-9 by validating the Emoqol-100 against PHQ-9 as a reference standard and to derivate the Emoqol-100 against the BDST questionnaire.

**Method**

A retrospective audit was conducted over 22 months, from 25/11/2020 to 28/9/2022, at a General Practice clinic in south Auckland, New Zealand. Participants were
consecutive patients (n=160) seen by one of the authors, in whom low mood was a key issue and who were coming for fully funded Wellness visits. A patient requests a wellness visit for a thirty-minute consultation, usually for emotional distress. The clinic gets paid an increased fee for service by the Health System. Some of these patients were regular, and clinic colleagues referred others for a Focussed Acceptance and Commitment Therapy (FACT) Consultation. Patients were eligible for the audit if they had a recorded Emoqol-100 score and a PHQ-9 score or BDST questionnaire administered at the same visit. These were assessed during the visit by the GP. The Emoqol-100 was the index test, and the BDST questionnaire or PHQ-9, was the reference standard. The order of doing the Emoqol-100, the BDST questionnaire and the PHQ-9 were variable, but the Emoqol-100 was generally done first as it was the quickest to complete and on some occasions, the only one done. The reference tests (PHQ-9 and BDST) were not given blindly to the patients. While the clinician had other information, such as medication and medical history, this did not alter the administration of the Emoqol-100 test, BDST or the PHQ-9. Only patients with reasonably good English language abilities were asked the Emoqol-100.

The analysis was done according to the method of (Guyatt, 2002) for calculating likelihood ratios. For each level of the Emoqol-100, the true positive number is divided by the total depressed and the false positive divided by the total not depressed (TP/all depressed) / (FP/those not depressed). A likelihood ratio greater than 1 increases the post-test probability of the condition, while a likelihood ratio less than 1 decreases the post-test probability of the condition.

A recent meta-analysis reported that a PHQ-9 score of ≥10 is the level where the combination of sensitivity and specificity is maximised overall, and this was the cut-off used. For BDST, a score ≥6 is classified as depression and was used as a cut-off. The number of patients available determined the sample size during the study period. There was no public or patient involvement in this work. Audits (if anonymous and retrospective) do not require Ethics approval by the National Health Ethics Committee of New Zealand.
Results

Baseline characteristics

There were 160 patients and n=426 visit records of PHQ-9 and n=513 visit records of BDST. The findings are shown in Figures 1 and 2. The majority of the patients (62%) were women, and the median age was 35 (14-78) (Table 1). The distribution of ethnicities reflects the general population of the clinical study site reasonably well. The median score of Emoqol-100 was 55. The median for PHQ-9 was 15, which is equal to mild-moderate depression, and the median for BDST was 8, equivalent to moderate low mood. Low mood was present in 69% of the sample according to PHQ-9, and in 67% according to BDST. The practice has 5000 registered patients. There are seven GPs, one GP trainee, two nurse practitioners who work as GPs four nurses and two health care assistants. The practice is called a very low-cost access (VLCA) practice, meaning that more than half the patients are either of Maori or Pacific ethnicity or live in the most socioeconomically deprived quintile. Because it is a VLCA practice, the clinic gets more funding from the health system.

Emoqol-100 validation against PHQ-9

For patients with an Emoqol-100 score of ≤20, the LR is 10.05, with a positive PHQ-9 (≥10) as the reference standard (Table 3). The Emoqol-100 score 21-40 had an LR of 2.9, score 41-60 had an LR of 1.3, score 61-80 had an LR of 0.40 and a score of 81-100 had an LR of 0.2 (Table 3). Based on the PHQ-9 ≥10, the highest positive predictive value was 96% for a cut-point of ≤20 (Table 3). This means that a patient who scores ≤20 is 96% likely to have a PHQ-9 score of 10 or more, indicating a high probability of a low mood at that visit and a clinically significant increase from the average low mood of 69%.

Emoqol-100 derivation against Burns Depression Scale Today

For patients with an Emoqol-100 score of ≤20, the LR of low Mood is 25.2, with a positive BDST (≥6) as the reference standard (Table 4). The Emoqol-100 score 21-40 had an LR of 3.6, score 41-60 had an LR of 1.7, score 61-80 had an LR of 0.4 and score 81-100 had an LR of 0.09 (Table 4). As the Emoqol-100 score gets higher, the LR drops, and in the higher Emoqol-100 range, the patient is much less likely to suffer from a mood disorder. Based on the BDST ≥6, the highest positive predictive value was 98% for a cut-point of ≤20 (Table 4). This means that a patient who scores ≤20 is
98% likely to have a BDST score of 6 or more which is a clinically significant increase from the average low mood on the BDST of 67%. The same applies for an Emoqol-100 score of > 80 the post-test likelihood of this is a positive predictive values of 15% which is a clinically significant decrease from 67%.

**Emoqol-100 against PHQ-9 Burns Depression Scale Today as a receiver operating curve**
The Emoqol (continuous) as a predictor of the PHQ-9 ≥ 10 had an area under the curve (AUC =0.7698, 95% CI 0.72270 to 0.81698) (not shown). An AUC of 0.7 to 0.8 is considered to be good. For the Emoqol-100 (continuous) as a predictor of the BDST ≥ 6 the AUC was 0.8192, 95% CI 0.78078 to 0.85762 (Figure 3). An AUC 0.8 to 0.9 is considered excellent (the ROC curves were done with STATA v 17).

**Discussion**

**Summary**
We validated the Emoqol-100 against the PHQ-9 and found the results consistent with the previous derivation study.(12) The Emoqol-100 score in the low range is associated with a high PHQ-9 and a high BDST (high scores indicate low mood). The higher the Emoqol-100 score is, the lower the scores for PHQ-9 and BDST. The Emoqol-100 score in the low range is associated with an increased risk of low mood, according to both the PHQ-9 and the BDST. For an Emoqol-100 score of 0-20 the LR for PHQ-9 is 10.1 and for BDST 25.2, which is very high and suggests that the test will significantly change post-test probabilities from pre-tests. It is unusual to have such high LRs in clinical medicine.(19) In addition, the Emoqol-100 with scores 81-100 have very low LRs, and a high score (a negative result) could be used to rule out low mood.

**Strengths and limitations**
This is a clinical audit and, as such, is a pragmatic study applied to consecutive patients with known or suspected depression. The advantage of an audit is that it is possible to measure inventories whilst in use in regular clinical practice, without the confounding issues of consent and information sheets which can cause a selection bias in planned research. The Emoqol-100 has been validated and derivated in the clinical setting where it is intended to be used. The prevalence of low mood was significantly higher than in a consecutive series of patients seen in a usual general practice setting. This sample of included patients who were seen by one of the authors for extra mental
health care. However, this should not affect the properties of the test but rather the interpretation of a negative test since a diagnosis is hard to rule out in a highly prevalent condition. This study has several limitations. The clinical audit only applies to one practitioner, and it was impossible to blind the measurement of the reference standard PHQ-9 or BDST. The PHQ-9 is becoming the common tool for primary care assessments of depression.\(^{(20)}\) The Emoqol-100 measures the mood right now and that can be considered both a weakness and a strength. It may be less valid than tools measuring more extended periods of low mood, but potentially more precise and without recall bias and faster to complete.

**Comparison with existing literature**

Low mood was present in 69\% (PHQ-9), respectively 67\% (BDST) of the cohort. This is a very high prevalence, and thereby pre-test probability is high compared to the average primary care prevalence of mood disorders of approximately 12.9\%.\(^{(21)}\) The prevalence was expected to be high as the patients were selected for extra attention for their mental health. The prevalence should not affect the LR results since it is a function of the sensitivity and specificity.\(^{(22)}\) Diagnostic properties like sensitivity and specificity are not influenced by changes in disease prevalence itself, whereas positive and negative predictive values directly depend on prevalence.\(^{(23)}\) The Emoqol-100 is not intended as a screening tool but needs to be evaluated in a screening setting.

The PHQ-9 takes around five minutes to complete and has been identified as a useful screening tool for depression.\(^{(4, 24)}\) However, using tools that take several minutes to apply and interpret and thus be impractical in a standard primary care setting, where consultations often are very short. Emoqol-100 is feasible, fast, easy and helpful in monitoring both severity and for following patients over time.

**Implications for research and practice**

The Emoqol-100 is the briefest of all mental health tools. It takes 15 seconds to verbally administer and interpret and has the advantage of patients not having to read. We speculate that by verbally asking the Emoqol-100 rather than doing a written questionnaire the opportunity for the patient to be seen, heard and understood is increased thereby enhancing beneficial patient-centred communication.\(^{(25, 26)}\) The
Emoqol-100 may also be helpful for monitoring mood over time since it specifically assesses the mood right now there should be no recall bias.

There are significant clinical changes with the high and low likelihood ratios. For example, with the BDST the pre-test probability is 67% and for a LR of $\leq 20$ the post-test probability of a low mood is 98%. For a LR $\geq 80$ the post-test probability is 15%. These are clinically useful changes in condition probability. It has the potential to be useful in clinical situations where low mood is suspected and time is limited, and a rapid assessment of mood would be clinically useful.

We suggest that the Emoqol-100 will be a valuable tool for clinicians in other clinical situations when there is not enough time to do a longer inventory, such as when emotional issues are raised near the end of a consultation or during a ward round. A score of 60 or below would indicate the need for further help and a longer appropriate intervention made at that visit or a subsequent appointment. For those scoring above 60, the Whooley two questions on the PHQ-2 about feeling depressed or losing pleasure during the last month could be asked; negative answers would most certainly rule out depression. (27)

Future research could systematically interview primary care clinicians and patients to ask how they feel about being asked these questions and what they think are the strengths and limitations. Our anecdotal experience of teaching the Emoqol-100 to groups of primary care clinicians is that they are highly positive about them. They realise that if the Emoqol-100 is inconsistent with the clinical picture or history, they can do a PHQ-9 or a BDST longer questionnaire.

For clinical practice, the US Preventive Services Task Force (2016) recommend "Commonly used depression screening instruments include the Patient Health Questionnaire (PHQ) in various forms and the Hospital Anxiety and Depression Scales in adults; the Geriatric Depression Scale in older adults, and the Edinburgh Postnatal Depression Scale (EPDS) in postpartum and pregnant women.(28) All positive screening results should lead to an additional assessment that considers the severity of depression and comorbid psychological problems (e.g., anxiety, panic attacks, or substance abuse), alternate diagnoses, and medical conditions."
Conclusions
The Emoqol-100 is a case-finding tool and has now been validated against the PHQ-9 and derived against the BDST. It could be used when the pre-test probability of a low mood is higher than usual, as in patients complaining of possible mood-related symptoms such as sleep difficulties, pain and fatigue. (29) The likelihood ratios at very high and very low Emoqol-100 scores are robust, creating clinically significant changes in post-test probabilities which is what is required in diagnostic tests. The Emoqol-100 is the briefest of all mental health tools we know. It and the BDST are the only inventories to measure how the patient is feeling today. The Emoqol-100 has the advantage of patients not having to read or interpret a written or verbally asked questionnaire. It takes only seconds to administer, which could be highly useful in case-finding in primary care. Further research is required using the Emoqol-100 in different clinical settings such as screening in primary care and clinical use in secondary care, by other clinicians, and as a tool for monitoring the ongoing mood state.

Disclosure statement
The authors declare no potential conflicts of interest.

Data availability
Anonymous data is available upon reasonable request.
Acknowledgement: Thanks to Prof Tim Kenealy for the ROC curves
A copy of the BDST is on supplementary tabel
Funding: There was no external funding for this project.
References

19. Kennedy KS, S. Cifu, A. Diagnostic Reasoning and the Accuracy of Findings in the History and Physical Examination. Forthcoming 2022. use the web address and format for an online reference
28. US preventive services task force recommendation on depression. https://www.uspreventiveservicestaskforce.org/uspstf/topic_search_results?category%5B%5D=20&searchterm=depression+
Figure 1. Flow diagram for records of reference standard PHQ-9 and Emoqol score, n=426 visits

- Distressed patients seen in time period n=160
- Eligible patients n=160
- n=426 visits with Emoqol and PHQ-9

- PHQ $\geq 10$ = depressed
  - N= 294 tests
  - True positive if Emoqol cut-off 20 n=45
  - False negative if Emoqol cut-off 20 n=249

- PHQ $< 10$ = not depressed
  - N= 132 tests
  - False positive if Emoqol cut-off 20 n=2
  - True Negative if Emoqol cut-off 20 N=130

Eligible patients n=160

Visits n=532

Distressed patients seen in time period n=160
**Figure 2.** Flow diagram for records of reference standard Burns D and Emoqol score, n=513 visits

- Distressed patients seen in time period n=160
- Eligible patients n=160
- n=513 visits with Emoqol and Burns D

- Burns D ≥ 6 = depressed
  - N= 345 tests
  - True positive if Emoqol cut-off 20 n=52
  - False negative if Emoqol cut-off 20 n=293

- Burns D < 6 = not depressed
  - N= 168 tests
  - False positive if Emoqol cut-off 20 n=1
  - True negative if Emoqol cut-off 20 n=167

Eligible patients n=160
Visits n=532
Distressed patients seen in time period n=160
Figure 3 (below). Emoqol (continuous) as predictor of Burns D >=6.
(AUC 0.8192, 95% CI 0.78078, 0.85762)
Table 1. Baseline measures n= 160 patients, Burns Depression Scale Today n=523, PHQ-9 questionnaires n= 434

<table>
<thead>
<tr>
<th>Age – median (SD)</th>
<th>35 years (±15.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female: 100 (62%)</td>
<td>Male: 60 (38%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>European: 79 (49%)</td>
<td>Māori: 38 (24%)</td>
</tr>
<tr>
<td>Indian: 13 (8%)</td>
<td>Samoan: 8 (5%)</td>
</tr>
<tr>
<td>*Other: 22 (14%)</td>
<td></td>
</tr>
<tr>
<td>BDST</td>
<td></td>
</tr>
<tr>
<td>Range (20-0)</td>
<td>Median: 8</td>
</tr>
<tr>
<td>PHQ-9</td>
<td>Range: (27-0)</td>
</tr>
<tr>
<td></td>
<td>Median: 14</td>
</tr>
<tr>
<td>Emoqol score</td>
<td>Range: (0-100)</td>
</tr>
<tr>
<td></td>
<td>Median: 55</td>
</tr>
</tbody>
</table>

PHQ-9 = 9-item Patient Health Questionnaire  
BDST = Burns Depression Scale Today  
*Other: Asian, Chinese, Cook Island Māori, Latin American, Middle east, Fijian, Tongan, Niuean

Table 2. Emoqol 100 versus Patient Health questionnaire PHQ-9 and BDST

<table>
<thead>
<tr>
<th>Emoqol 0-100 range</th>
<th>PHQ-9 median range 0 to 27</th>
<th>BDST median range 0 to 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>21-40</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>41-60</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>61-80</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>81-100</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Range</td>
<td>Likelihood ratio +ve</td>
<td>Positive Predictive Value</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>0-20</td>
<td>10.1</td>
<td>96%</td>
</tr>
<tr>
<td>21-40</td>
<td>2.9</td>
<td>87%</td>
</tr>
<tr>
<td>41-60</td>
<td>1.3</td>
<td>74.5%</td>
</tr>
<tr>
<td>61-80</td>
<td>0.4</td>
<td>47.4%</td>
</tr>
<tr>
<td>81-100</td>
<td>0.2</td>
<td>15%</td>
</tr>
</tbody>
</table>

PHQ-9 = 9-item Patient Health Questionnaire
Likelihood ratio positive = sensitivity/(1-specificity)
\(a= \) also known as the post-test likelihood of a positive test
Likelihood ratio positive = sensitivity/(1-specificity) (these scores are > 1.0)
Likelihood ratio negative = 1-sensitivity/specificity) (these scores are < 1.0)
Table 4. Derivation assessment of Emoqol 0-100 with reference standard Burns Depression Scale Today (BDST)

<table>
<thead>
<tr>
<th>Range</th>
<th>Likelihood ratio +ve</th>
<th>Positive Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20</td>
<td>25.2</td>
<td>98%</td>
</tr>
<tr>
<td>21-40</td>
<td>3.6</td>
<td>88%</td>
</tr>
<tr>
<td>41-60</td>
<td>1.7</td>
<td>78%</td>
</tr>
<tr>
<td>61-80</td>
<td>0.35</td>
<td>41%</td>
</tr>
<tr>
<td>81-100</td>
<td>0.09</td>
<td>15%</td>
</tr>
</tbody>
</table>