SCREENING FOR DEPRESSION IN OLDER ADULTS ON AN ACUTE MEDICAL WARD: THE VALIDITY OF NICE GUIDANCE ON USING TWO QUESTIONS.

Collins Esiwe[1], Sarah Baillon[2], Aniruddha Rajkonwar[3], James Lindesay[2], Nelson Lo[4], Michael Dennis[5]

[1] Lincolnshire Partnership NHS Foundation Trust, Sleaford, Lincolnshire, UK

[2] Department of Health Sciences, University of Leicester, Leicester, UK

[3] Tees, Esk and Wear Valley NHS Trust, County Durham, UK

[4] Department of Geriatric Medicine, University Hospitals of Leicester NHS Trust, Leicester, UK

[5] College of Medicine, Swansea University, Wales, UK

Corresponding Author:

Sarah Baillon, Research Fellow, Department of Health Sciences, University of Leicester, 22-28 Princess Road West, Leicester, LE1 6TP

Email: sfb5@leicester.ac.uk

Tel. 0116 2588161, Fax. 0116 2584078

ABSTRACT

Background: Depression is common in older people in general hospital settings and associated with poor outcomes. This study aimed to evaluate the validity of two screening questions recommended by the UK National Institute for Health and Clinical Excellence (NICE).

Methods: 118 patients aged over 65 years, admitted to acute medical wards at a teaching hospital, were interviewed in a standardised manner using relevant sections of the Present State Examination – Schedules for Clinical Assessment in Neuropsychiatry to identify depression according to ICD-10 criteria. Subsequently participants completed the two depression screening questions and the 15 item version of the Geriatric Depression Scale (GDS-15).

Results: A threshold of one or more positive responses to the two NICE depression screening questions gave a sensitivity of 100%, specificity of 71%, positive predictive value (PPV) 49%, and negative predictive value (NPV) 100%. The GDS-15 optimal cut-off was 6/7 with a sensitivity of 80%, specificity of 86%, PPV 62%, and NPV 94%. A two-stage screening process utilising the NICE two questions followed by the GDS-15 with these cut-offs gave a sensitivity of 80%, specificity of 91%, PPV 71%, and NPV 94%.

Conclusion: The two depression questions perform well as an initial screening process for non-cognitively impaired older people in the acute medical setting. A positive response to either question would indicate further assessment is required by a clinician competent in diagnosing depression in this population, or the possible use of a more detailed instrument such as the GDS-15 to reduce the number of false positive cases.

INTRODUCTION

Depression in older people in the general hospital setting is very common, with a prevalence of 29% [1]. As well as the suffering associated with the disorder, the presence of depression is associated with a variety of adverse outcomes with increased mortality and morbidity, reduced compliance with treatment and rehabilitation, prolonged length of stay and increased likelihood of care home placement [1]. However, depression in older people in general hospitals is frequently undetected or poorly managed [2]. Reasons for this are complex but could include: a lack of skill, experience, time and confidence of more junior medical staff to make a diagnosis; difficulty interpreting the relevance of somatic symptoms; and the denial of low mood by older people.

Because of the potential benefits of appropriate treatment and support, the U.K. National Institute for Health and Clinical Excellence [3,4] recommend screening for depression in high risk patient groups and adults with chronic physical health problems with the use of the two screening questions [5]:

- *During the past month, have you often been bothered by feeling down, depressed or hopeless?*

- *During the past month, have you often been bothered by having little interest or pleasure in doing things?*

These screening questions have been widely adopted in clinical settings [6], although the evidence for use of the two screening questions in older people in the acute medical setting is limited. A recent systematic review of depression screening instruments in this clinical setting [7] identified only one study that utilised the two questions [8]. The Geriatric Depression Scale [9] was the most validated instrument identified in this review.

The objectives of our study were to investigate the effectiveness of the two screening questions recommended by NICE, and compare their performance with an instrument with established validity, namely the 15 item version of the Geriatric Depression Scale.

METHODS

Participants were recruited from four acute medical wards at the Leicester General Hospital, in University Hospitals of Leicester NHS Trust between October 2006 and June 2009. Patients were considered eligible for participation if they were: aged 65 years or over; medically fit to be interviewed; English-speaking; and considered to have capacity to give informed consent. All potential participants were identified by the ward medical team and approached to see if they were interested in participating in the research - this could have occurred at any point during their hospital admission episode.

Those who expressed an interest in the study were given an information sheet, and if they subsequently indicated that they were willing to participate they were then visited on the ward by a research doctor (CE or AR) who took written informed consent. First, Section 21 of the Present State Examination – Schedules for Clinical Assessment in Neuropsychiatry (SCAN) [10] which incorporates the Mini-Mental State Examination [11] was completed with the participant. If they scored less than 24 on the MMSE they were excluded from the remainder of the study - this cut-off is widely used to exclude patients with dementia [12, 13]. Participants who scored 24 or more on the MMSE then completed the sections of the SCAN interview pertaining to depression (sections 7, 8 and 9). A diagnosis of ICD-10 depressive episode was obtained by running diagnostic algorithms in the SCAN software; the gold standard in the study. The research doctors were Old Age psychiatrists who had received formal training in the SCAN interview technique.

Within 5 days of the initial interview participants were then visited by another researcher (SB), who was blind to the outcome of the first assessment, and completed the two NICE screening questions and the GDS-15 [14]. Participants were given the choice of completing the questionnaire themselves, reading the items and indicating their chosen response verbally to the researcher, or having the items read to them verbatim.

The project was approved by the Leicestershire, Northamptonshire and Rutland Research Ethics Committee (Ref 06/Q2501/39).

Sensitivity (Se), specificity (Sp), positive and negative predictive value (PPV, NPV), positive and negative likelihood ratio (LR+, LR-) and the diagnostic odds ratio (DOR) were calculated. Receiver operating characteristic (ROC) curve analyses were carried out using the methods described by Altman [15] and the areas under the curves (AUC) were compared using the methods described by Hanley and McNeil [16]. Analyses were performed using SPSS v18.

RESULTS

A total of 139 participants consented to take part in the study. Out of these, 6 scored less than 24 on the MMSE and were excluded, 2 were excluded as they were found to be younger than 65 years, one consented but was discharged before the SCAN interview was completed, 10 completed the SCAN interview but declined to complete the screening instruments when visited by the researcher, 1 completed the SCAN interview and was then discharged but could not be contacted by the researcher, and 1 completed the SCAN interview and was transferred to an acute psychiatric assessment ward for treatment of severe depression so the researcher was no longer blind to their depression status. Therefore 118 participants completed both the SCAN and the depression screening instruments and were included in the analysis. One participant did not complete the GDS-15 but completed all other assessments. Where available, age (n=15), MMSE (n=10 without cognitively impaired exclusions) and gender (n=16) of non-completers were compared to the 118 participants using the *t* test for continuous variables and chi-squared test for proportions (Fisher's Exact test). Non completers had similar age (*P* = 0.18) and MMSE scores to completers (*P* = 0.83) though had a higher proportion of males (*P* = 0.05).

Twenty six participants were diagnosed with depressive disorder according to the ICD-10 criteria (see Table 1 for characteristics of all participants).

A total of 24 (20.3%) participants responded positively to one of the two NICE screening questions, and 29 (24.6%) responded positively to both.   The optimal threshold for the identification of depression was 1 – i.e. responding positively to one or both of the screening questions (Table 2). The performance characteristics of the GDS-15 in this sample are shown in Table 3. If the recommended cut-off score of 5/6 [14] is applied the sensitivity was 80% and specificity of 85% and a PPV of 60%. The cut-off 6/7 gave a slightly better combination of sensitivity of 80%, specificity of 86% and a PPV of 62%.

The AUC for the identification of depression using the NICE questions, calculated from the ROC curve, was 0.92 (p<0.001, 95% confidence interval 0.87 to 0.97).   The AUC for the GDS-15 was 0.88 (p<0.001, 95% confidence interval 0.81 to 0.96).

One component of NICE guidelines [3] was a suggestion of using a two-stage screening process in which the two questions are the initial screen, followed by a second screening measure for those who respond positively to either or both of the questions. We used data from this study to simulate this two-stage screening process. 52/117 participants responded positively to one or both of the two NICE screening questions (resulting in 27 false-positives). Using the 6/7 threshold on the GDS-15 after initial screening with the two questions correctly identified 20/25 participants with depression (80% sensitivity) with an improved overall specificity (91%), a PPV of 71%, NPV 94%, LR+ 9.20, and LR- 0.22. Using the standard 5/6 threshold on the GDS-15 produced the same results. Using this two-staged process would reduce the false-positives from 27 to 8 cases but this would be at the cost of 'missing' 5 participants with depression who were incorrectly identified as non-depressed by the GDS-15 (but just one case of depression of moderate severity).

DISCUSSION

The primary objective of our study was to examine the validity of the two depression screening questions recommended by NICE for older people in the acute medical setting. The two questions appeared a useful measure in identifying patients with a depressive episode – a 'yes' response to either of the two questions identified all cases of depression (100% sensitivity). However, this detection rate at the recommended 0/1 cut-off comes at a cost, with only 71% of non-depressed identified by two 'no' responses (71% sensitivity), a positive likelihood ratio of less than 5 (3.4) reflecting poor accuracy, and a positive predictive value of 0.49 indicating half the patients who respond yes to one or both questions are false-positives. The performance of the two screening questions in our study was better however than Blank and colleagues [8] study identifying major depression in elderly people in the hospital setting. In their study using the same questions (PRIME-MD 2) the sensitivity was 92%, but specificity was only 54%, with a LR+ of 2. The major depression criteria are stricter than ICD-10 depression; this may account for the different performance of the two screening questions in our study compared to Blank et al who also screened all eligible patients rather than relying on ward staff to identify potential participants.

The GDS-15 demonstrated a good performance in our study, with a cut-off of 6/7 marginally better than 5/6. At a 6/7 cut-off the sensitivity was 80%, specificity 86%, LR+ 6.13, LR – 0.25, and a Diagnostic Odds Ratio of 26.67 indicating the instrument would be useful in clinical practice. The GDS is the only instrument that has been adequately examined in this clinical context [7]; in a pooled analysis the GDS-15 with a 5/6 cut-off had a sensitivity of 79%, specificity of 77%, LR+ of 3.4, and LR – 0.29. The GDS-15 also performed better in our study than Cullum *et al.* [17] who also found an optimum cut-off of 6/7 but with a sensitivity of 74%, specificity of 81%, LR+ 3.98, LR- 0.32, DOR 12.62.

NICE guidance on depression in adults with a chronic physical health problem suggest that a clinician who is competent in a mental health assessment ask further questions to determine the presence of

mental disorder, and if not competent, refer to a competent professional [4]. In the case of a patient in the general hospital setting such a professional would be a liaison psychiatrist, or in some circumstances another member of the treating team. With nearly half the patients in our study responding positively to one or both questions (53/118; 45%), of whom half again are false positives unless the assessing clinician is competent in diagnosing depression in this age group this would mean substantial numbers referred on to liaison services. There would be substantial difficulties in many areas; liaison services are frequently rudimentary or under-resourced with issues concerning availability [18] despite an increasing evidence base to support their development [19, 20]. The NICE updated guidance for depression [3] suggests the next possible step for an 'at risk' adult who has answered positively to one or both screening questions could be a further assessment using a recognised screening instrument. This two stage process was adopted by the 'Identifying depression in hospital' target in Wales [6]. This programme suggested using the Patient Health Questionnaire (PHQ-9). The PHQ-9 however has not been validated in older people in the acute medical setting, whereas the GDS-30 and GDS-15 have [7]. In our study we were able to examine the use of a two stage screening process involving the two NICE screening questions followed by the GDS-15. This method identified 20/25 participants with depression (80% sensitivity) with an improved overall specificity (91%), a PPV of 71%, NPV 94%, LR+ 9.20, and LR- 0.22. Although the false positives are reduced to 8, 5 cases of depression are missed. However, only one case of depression of moderate severity was missed by this process; this is of particular pertinence because it is moderate and severe depression that responds best to antidepressant treatment [21]. The stepped care approach adopted by NICE [3,4] emphasises the important role of low-intensity psychosocial interventions, education, and support for those with sub-threshold or mild depression so it is reasonable to suggest that all patients screening positive on the NICE questions warrant further assessment and support during their in-patient stay and on discharge.

Despite NICE guidance, there remains a debate concerning the efficacy of screening programmes [22, 23], and evidence for screening and treating depression in older people in medical in-patient units is

mixed. U.K studies have shown benefits of a collaborative care approach to ensure effective treatment and appropriate follow-up with mental health liaison services for older patients with depression in general hospitals [24, 25], whereas Cole and colleagues [26] in Canada failed to show that depression screening and multidisciplinary care improved outcomes and depression syndromes remained stable after discharge regardless of the effects of antidepressants [27].

Limitations

Although comparable with most studies examining depression screening instruments in this clinical population, the sample is relatively small. Because of the opportunistic use of ward medical teams to identify potential participants we cannot be certain that our sample is truly representative and there is the possibility that staff may be more likely to approach those they think are depressed. However, although our non-completers were more likely to be male, the age and gender characteristics of the participants in our study are similar to studies of older people with unscheduled general hospital admissions in England [28, 29]. Additionally, as patients with cognitive impairment were excluded it is not possible to generalise the findings to people with dementia. As this is the first time a two-stage screening process has been examined, in particular incorporating the two NICE questions, and our analysis was based on a sub-sample of participants (52/117) further similar investigations are required. Another limitation is the absence of any cases of severe depressive episode, the reasons for this could be the presence of co-existing cognitive impairment related to depression, lack of capacity of such patients, or unwillingness to volunteer. It may also have been preferable to have a balanced order of study assessments to avoid the possibility of any priming effect of the SCAN on the subsequent assessments, or to have completed the screening scales followed by the SCAN interview in order to replicate the order of presentation in practice but for practical reasons neither were possible. Additionally, although the SCAN interview and screening instruments were administered a maximum of five days apart, we cannot fully discount the possibility of a delirium that may have been present initially and resolved or developed subsequently; depressive symptoms can frequently be a component of a delirium episode [30].

Conclusions

Our study is the first to specifically examine the two depression screening questions recommended by NICE in older people in the acute medical setting and generally supports the guidance of this being the first step in screening for clinically relevant depression. However, there are a high number of false positives and the next step in confirming diagnosis will depend upon other factors, in particular whether the clinician is sufficiently skilled to inquire further about the presence of other depressive symptoms, and the accessibility and availability of liaison mental health services. A second stage to screening using a validated instrument such as the GDS-15 may be useful in some circumstances. Further research is required, in particular: using random samples; examining acceptability of depression screening to patients and carers; and randomized controlled trials to evaluate the effectiveness of screening for depression in this setting.

Key points

- Depression is common in older people in acute medical in-patient units.
- The 2 questions advocated by NICE for screening for depression in this setting performed well in cognitively intact patients.
- A 2-stage screening process incorporating the two questions followed by GDS-15 would reduce the number of false positive cases.

CONFLICTS OF INTEREST

None declared.

REFERENCES

1. Royal College of Psychiatrists. Who cares wins. Improving the outcome for older people admitted to the general hospital: Guidelines for the development of Liaison Mental Health Services for Older People. London 2005. http://www.rcpsych.ac.uk/PDF/WhoCaresWins.pdf.

2. Philpot M, Lyons D, Reynolds T. Liaison old age psychiatry in the general hospital. In: Jacoby R & Oppenheimer C (Eds.) Psychiatry in the Elderly, 3rd edition, pp401-419, Oxford: Oxford University Press, 2002.

3. National Institute for Health and Clinical Excellence. Depression in Adults (update). Depression: the treatment and management of depression in adults. National Clinical Practice Guideline 90. London: National Institute for Health and Clinical Excellence, 2009.

4. National Institute For Health and Clinical Excellence Depression in adults with a physical health problem: treatment and management. London NICE (www.nice.org/CG91)2009.

5. Whooley M, Avins A, Miranda J. et al. Case-finding instruments for depression. Two questions as good as many. Journal of Internal Medicine, 1997; 12:439-445.

6. NHS Wales 2012 1000 lives + Mental Health - Identifying Depression in Hospital Settings http://www.1000livesplus.wales.nhs.uk/mh-depression

7. Dennis M**,** Coffey J, Kadri A. 2012 Depression in older people in the general hospital: a systematic review of screening instruments Age & Ageing, 41, 148-154

8. Blank K, Gruman C, Robison J. Case-finding for depression in elderly people: balancing ease of administration with validity in varied treatment settings. Journal of Gerontology, 2004; 59; 378-384

9. Yesavage J, Brink T, Rose T et al., Development and validation of a geriatric depression scale: a preliminary report. Journal of Psychiatric Research, 1983;17: 37-49

10. World Health Organisation, *Schedules for Clinical Assessment in Neuropsychiatry Interview, Version 2.1*, 1999, World Health Organisation: Geneva.

11. Folstein, M., Folstein, S., McHugh, P. , *Mini-mental State – a practical method for grading the cognitive state of patients for the clinician.* Journal of Psychiatric Research, 1975;12: 189-198.

12. Tombaugh T, McIntyre N. The mini-mental state examination: a comprehensive review. Journal of the American Geriatric Society, 1992; 40: 922-935

13. Folstein M, Folstein S, McHugh P, Fanjiang G. Mini-Mental State Examination users guide. Odessa: Florida; 2001.

14. Sheikh J, Yesavage JA. Geriatric Depression Scale: recent evidence and development of a shorter version. In: Brink TL, editor. Clinical Gerontology: A guide to assessment and intervention. New York: Haworth; 1986. p. 165-73.

15. Altman, D., *Practical statistics for medical research* 1999: Chapman & Hall/CRC.

16.Hanley, J.A., McNeil, B.J., *A method of comparing the Areas under the Receiver Operating Characteristic curves derived from the same cases.* Radiology, 1983. **148**: p. 839-843.

17. Cullum S, Tucker S, Todd C, Brayne C. Screening for depression in older medical inpatients. International Journal of Geriatric Psychiatry 2006; 21: 469-476

18. NHS Confederation Mental Health Network (2012) Liaison psychiatry – the way ahead. Nov, 249.

19. Parsonage M, Fossey M., Tutty C. *Liaison psychiatry in the modern NHS.* Centre for Mental Health. www.centreformentalhealth.org.uk/publications 2012

20. Tadros G, Salama R, Kingston P et al. Impact of an integrated rapid response psychiatric liaison team on quality improvement and cost savings: the Birmingham RAID model. The Psychiatrist, 2013; 37: 4-10.

21. Dowrick C, Frances A. Medicalising and medicating unhappiness. BMJ, 2013; 347, f7140

22. Gilbody S, House AO, Sheldon TA. Screening and case finding instruments for depression. Cochrane Database Syst Rev. 2005:CD002792. [PMID: 16235301].

23. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. CMAJ 2008; 178(8):997-1003.

24. Baldwin R, Pratt H, Goring H et al. Does a nurse led mental health liaison service for older people reduce psychiatric morbidity in acute general medical wards? A randomised controlled trial. Age and Ageing, 2004; 33:472-478

25. Cullum S, Tucker S, Todd C, Brayne C. (2007) Effectiveness of liaison psychiatric nursing in older medical inpatients with depression: a randomised control trial. Age and Ageing, 2007;36:436-442.

26. Cole M, McCusker J, Elie M, Dendukuri N, Latimer E, Belzile E. Systematic detection and multidisciplinary care of depression in older medical inpatients: a randomized trial. CMAJ, 2006; 174: 38-44

27. McCusker J, Cole M, Ciampi A, Latimer E, Windholz S, Belzile E. Twelve-month course of depressive symptoms in older medical inpatients. International Journal of Geriatric Psychiatry, 2007; 22: 411-417

28. Walsh B, Roberts H, Nicholls P. Features and outcomes of unplanned hospital admissions of older people due to ill-defined (R-coded) conditions: Retrospective analysis of hopital admissions data in England. BMC Geriatrics, 2011; 11:62

29. Glover A, Bradshaw L, Watson N, et al. Diagnosis, problems and healthcare interventions amongst older people with an unscheduled hospital admission who have concurrent mental health problems: a prevalence study. BMC Geriatrics, 2014; 14:43.

30. Farrell K, Ganzini L. Misdiagnosing delirium as depression in medically ill elderly patients. Archives of Internal Medicine, 1995; 155: 2459-2464.

Table 1. Characteristics of the study participants (N = 118)

Gender	Male : Female	41 (34.7%) : 77 (65.3%)
Age	Median (IQR)	82.0 yrs (76.8-86.0)
MMSE Score	Median (IQR)	27.0 (26.0-28.0)
ICD-10 diagnosis, N (%)	No depression	92 (78%)
	Mild depressive episode	16 (13.5%)
	Moderate depressive episode	10 (8.5%)
	Severe depressive episode	0

Table 2. Operating characteristics of the two NICE screening questions for the identification of ICD-10 depressive episode

Cut-off	Se (95% CI)	Sp (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+	LR-	Diagnostic Odds Ratio
0/1	1.00	0.71	0.49	1.00	3.41	0	--
	(0.83-1.0)	(0.60-0.79)	(0.35-0.62)	(0.93-1.0)			
1/2	0.77	0.90	0.69	0.93	7.86	0.26	30.74
	(0.55-0.90)	(0.81-0.95)	(0.49-0.84)	(0.85-0.97)			

Table 3. Operating characteristics of the GDS-15 for the identification of ICD-10 depression (N=117)

Cut-off	Se (95% CI)	Sp (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+	LR-	Diagnostic Odds Ratio
2/3	0.96 (0.77-0.99)	0.49 (0.38-0.59)	0.33 (0.23-0.46)	0.97 (0.87-0.99)	1.88	0.08	22.98
3/4	0.88 (0.67-0.96)	0.59 (0.49-0.69)	0.37 (0.25-0.50)	0.94 (0.84-0.98)	2.19	0.20	10.90
4/5	0.84 (0.63-0.94)	0.75 (0.64-0.83)	0.47 (0.32-0.63)	0.94 (0.85-0.98)	3.36	0.21	15.75
5/6	0.80 (0.58-0.92)	0.85 (0.76-0.91)	0.60 (0.42-0.76)	0.94 (0.86-0.97)	5.66	0.23	24.31
6/7	0.80 (0.58-0.92)	0.86 (0.77-0.92)	0.62 (0.43-0.78)	0.94 (0.86-0.97)	6.13	0.23	26.67
7/8	0.64 (0.42-0.81)	0.91 (0.83-0.95)	0.66 (0.44-0.83)	0.90 (0.81-0.95)	7.36	0.39	18.67
8/9	0.60 (0.38-0.78)	0.93 (0.85-0.97)	0.71 (0.47-0.87)	0.89 (0.81-0.94)	9.20	0.43	21.50
9/10	0.48 (0.28-0.68)	0.94 (0.87-0.97)	0.70 (0.44-0.86)	0.87 (0.78-0.92)	8.83	0.55	16.06