Screening for dementia in general hospital inpatients: a systematic review and meta-analysis of available instruments

THOMAS A. JACKSON1,2, SYEDA HUMA NAQVI2, BART SHEEHAN3

1School of Immunity and Infection, University of Birmingham, Birmingham, UK
2Department of Geriatric Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
3Psychological Medicine, John Radcliffe Hospital, Oxford, UK

Address correspondence to: T. A. Jackson. Tel: 0121 371 3264; Fax: 0121 371 3203. Email: thomasjackson@doctors.org.uk; t.jackson@bham.ac.uk

Abstract

Objective: Dementia is common and often undiagnosed. Improving rates of diagnosis has become a key part of current dementia guidelines. Older people admitted to hospital are a potential target population for screening for dementia. The objective was to report whether instruments advocated in screening for dementia had been validated in hospital inpatients and to make recommendations on evidence-based screening for dementia in this population.

Design: a systematic review was performed by an initial electronic database search using three key search criteria. Studies were then selected in a systematic fashion using specific predetermined criteria. Pooled meta-analysis was performed. Inclusion criteria were studies where the study group were inpatients in general hospitals, including a clearly defined group of older people (60 or older), they used a recognised screening instrument compared with a reference standard, and included at least 10 cases of dementia. Demographic data as well as sensitivity and specificity were recorded from the selected studies.

Results: in total nine studies describing validation of six discreet instruments satisfied all our criteria and we were able to perform meta-analysis with one instrument, the Abbreviated Mental Test Score (AMTS). With a cut-off of <7, pooled analysis of the AMTS showed a sensitivity of 81%, a specificity of 84% and an area under the curve (AUC) of 0.88.

Conclusion: a small number of instruments have been validated for screening for dementia in general hospital. Understanding strengths and weaknesses of currently available instruments allows informed decisions about screening in this setting.

Keywords: dementia, screening, general hospital, older people, systematic review

Introduction

Dementia affects 36 million people worldwide, with numbers expected to double every 20 years to 66 million by 2030 [1]. In the UK 700 000 people are affected, but it is estimated that only a third of people with dementia currently have been diagnosed [2]. Improving diagnosis rates is a key aspect of the World Alzheimer’s Report of 2011[3]. Despite a drive in the neurosciences for biomarkers to detect early dementia [4] the ethics of when, or indeed whether to diagnose what is a progressive neurological disease remain difficult [5] and some take the view that diagnosis is of little benefit [6]. This may reflect a negative view among clinicians of what can be done therapeutically [7], and the case for general population screening for dementia remains controversial [8, 9].

There is potential value in the diagnosis of dementia throughout the disease course. Diagnosis allows access to the appropriate support services as well as drug treatment. There is clinical evidence demonstrating drug efficacy in early dementia [10, 11]. Earlier diagnosis may allow patients to make advance care decisions while still competent to do so. Diagnosis may also improve the quality of life for carers by allowing access to dementia-specific resources as well as providing an explanation for a persons altered mental state [12]. Economically, in the UK early diagnosis would mean increased costs up front, although savings made through reduced institutionalisation and
better care might result in overall costs savings and health benefits [13, 14]. Diagnosis is well accepted by patients with a survey of patients in a memory clinic showing that even patients with severe dementia would prefer to be told their diagnosis [15]. Although diagnosis has benefits it should also be appreciated that it can have major psychosocial effects, both positive and negative [16].

Dementia and cognitive impairment are common among older hospital patients and remain under diagnosed [17, 18], in common with other mental health problems in the same group [19]. In 2009 Sampson et al. [17] showed that 42% of unselected older medical inpatients had dementia: half of these had not previously been diagnosed with dementia, while mortality was much higher among those with dementia. The diagnosis of dementia in general hospitals is also complicated by the complex diagnostic challenge of concurrent delirium, with up to two-thirds of people with delirium having concurrent dementia, and dementia itself being a significant risk factor for the development of delirium [20]. The need to improve diagnosis of dementia in hospitals is long established [21]. Medical admission could therefore offer a timely opportunity to identify potential cases of dementia. Recognition of dementia also allows for improved care during the hospital admission. Improved care may include avoiding new medical events known to be more likely among inpatients with dementia [22], accessing dementia services, planning legal and capacity assessments and involving family in care decisions [23].

There are many screening instruments in current use and guidelines exist on which tests to use. These are often not restricted to hospital use and may not be validated for hospital use. Current guidelines include the National Institute for Health and Clinical Excellence (NICE), who in their UK guidance on recognition of dementia suggest using the Mini-Mental State Examination (MMSE) [24], the 6 item cognitive impairment test (6-CIT) [25], the general practitioner assessment of cognition (GPCOG) [26] and the 7-minute screen [27, 28]. Guidelines from the American Geriatrics Society recommend using the Mini-Cog assessment instrument for dementia (Mini-cog) [29] followed by the MMSE or Montreal Cognitive Assessment (MOCA) [30] if positive [31]. A further tool in common use is the Addenbrooke’s Cognitive Assessment [32]. The British Geriatrics Society suggests using the MMSE, the CLOX1, an executive clock drawing test (CLOX1) [33] and the Informant Questionnaire of Cognitive Decline in the Elderly (IQCODE) [34], in conjunction with a delirium screen, to identify dementia specifically in medical inpatients [35]. All these professional guidelines emphasise a two-stage approach; that is, detailed assessment after initial screening or clinical suspicion.

The aim of this review is to determine which of the instruments advocated for screening for dementia have been validated in older hospital inpatients and therefore inform decision-making for services.

Search strategy and selection criteria

An electronic database search of Embase, PsycINFO and MEDLINE was made for articles in English using search terms in the following three domains: dementia and cognitive impairment, diagnosis and screening tests, and thirdly general hospital inpatients. Supplementary data available in Age and Ageing online, Appendix 1 contain the full search strategy. The databases were accessed on 20 November 2012 and Embase was searched from 1947, PsycINFO from 1967 and MEDLINE from 1946. Only English language articles were accessed due to lack of resources to translate. The abstracts were then screened by two assessors (T.J., H.N.) independently based on the following criteria and the full texts then retrieved if they:

- Included patients studied during a hospital inpatient stay using a screening test for cognitive impairment or dementia. Patients studied in psychiatric wards, memory clinics and the community were excluded
- Included an age defined group of older people (60 years or older)

Published review articles on cognitive screening were also examined to identify any further studies [36–43]. An additional electronic search was done with each identified instrument as a key search term with the search terms for general hospital inpatients to identify any further validation studies. The reference sections of the selected papers were also studied as were relevant clinical guidelines.

The full texts of the selected studies were reviewed independently by three reviewers (T.J., H.N. and B.S.) against the following final inclusion criteria:

- The study groups are inpatients in general hospitals, and not in a psychiatric hospital, care home or the community.
- Studies include older people (60 years or older) as the main subject group or a clearly defined subgroup.
- The study uses a recognised screening instrument for cognitive impairment or dementia and this is compared with a ‘gold’ or reference standard. The reference standard was defined as the Diagnostic and Statistical Manual of Mental Disorders (DSM) versions III to IV [44], International Classification of Diseases – 10th edition (ICD-10) [45], the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association Alzheimer’s criteria (NINCDS-ADRDA) [46] or expert diagnosis following interview.
- At least 10 cases of dementia according to gold standard.

From the selected papers, the following data were recorded: patient group and sample size, mean age and proportion female, prevalence of dementia, test used and cut-off, comparator (gold standard), sensitivity and specificity. Any disagreements were decided by consensus.

Following previous convention, statistical analysis was done to produce a meta-analysis if there were three studies assessing the same test [47]. We used Meta-Disc version 1.4, a meta-analytical software package [48] to produce a pooled sensitivity, specificity, positive likelihood ratio (LR +), negative likelihood ratio (LR –), diagnostic odds ratio (DOR) and a summary receiver operating characteristics analysis (SROC). Likelihood ratios are a measurement of diagnostic
accuracy and say how likely a person with a condition is to have a positive test (LR+) or a person without the condition is to have a positive test (LR−). The diagnostic odds ratio is also a measurement of diagnostic accuracy independent of prevalence and represents the probability of the test being positive if a person has the disease relative to the odds of the test being positive if the person does not have the disease. Heterogeneity was measured by calculating $I^2$. The $I^2$ index describes the variation across the studies that due to significant heterogeneity rather than random chance.

Selected studies were then reviewed against the QUADAS-2 criteria to assess the study quality and risk of bias [49].

**Results**

The initial search returned 447 articles of which 18 were selected. Three further articles were identified by the review articles selected and a further three identified by searching by specific tests. See Figure 1 for details.

Of those 24, 9 studies met the inclusion criteria [50–58]. Six were excluded as the MMSE was used as the main comparator, six were excluded as the setting was not exclusively the general hospital and three were excluded as they did not involve a test to diagnose dementia. Six discrete instruments were investigated in the nine studies. The papers using a recognised gold standard diagnosis as comparator are shown in Table 1. The instruments studied were the Abbreviated Mental Test Score (AMTS), the Digit Span backwards test, the Time and Change Test, the Informant Questionnaire of Cognitive Decline in the Elderly short form (IQCODE), the Short Portable Mental Status Questionnaire (SPMSQ) and the MMSE. Across the studies sensitivity ranges from 73 to 100% and specificity ranges from 65 to 99%. All the studies involved medical inpatients and dementia prevalence varies from 6 to 52.1%.

Meta-analysis of the three studies comparing the AMTS (cut-off value of <7) with dementia as defined by DSM-IIIR was performed (Table 2) [52, 54, 55]. This showed an estimated prevalence of dementia of 7.8% with a sensitivity of 81%, specificity of 84%, a positive likelihood ratio of 5.05, a negative likelihood ratio of 0.23 and a diagnostic odds ratio of 22.37. Figure 2 shows a SROC for the three studies with an area under the curve (AUC) of 0.88. SROC derivation and details of the Q* are reported elsewhere [59].

Study quality and risk of bias assessed using QUADAS-2 is shown in Supplementary data are available in Age and Ageing online, Table S1. There are two main concerns about potential risk of bias. Five studies did not blind the assessors’ of the reference standard to the initial screening test result, and the index test threshold was not pre-specified in seven studies. All studies were otherwise well designed and there was no concern regarding the study applicability.

---

*Figure 1. Flow diagram showing the selection of studies for the systematic review and meta-analysis.*
Table 1. Data from studies using the reference standard

<table>
<thead>
<tr>
<th>Study and setting</th>
<th>Patient group, age and sample size</th>
<th>Mean age</th>
<th>%♀</th>
<th>Comparator</th>
<th>Prev (%)</th>
<th>Test used and cut-off</th>
<th>Sen</th>
<th>Sp</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Keefe et al. [50]</td>
<td>Medical inpatients Over 65 years, n = 169</td>
<td>76, Range 65–101</td>
<td>58</td>
<td>Interview and expert opinion</td>
<td>37</td>
<td>Any error in year with orientation</td>
<td>85</td>
<td>92</td>
</tr>
<tr>
<td>Leung et al. [51]</td>
<td>Medical inpatients Over 75 years, n = 144</td>
<td>80, no SD reported</td>
<td>46</td>
<td>DSM-IV</td>
<td>36</td>
<td>Any error in month with orientation</td>
<td>77</td>
<td>78</td>
</tr>
<tr>
<td>Antonelli Inclaz et al. [52]</td>
<td>Medical Inpatients n = 2808</td>
<td>71, ±0.27 SE</td>
<td>45</td>
<td>DSM-III-R</td>
<td>6.0</td>
<td>Digit span backwards Unable to complete 3 numbers</td>
<td>81</td>
<td>84</td>
</tr>
<tr>
<td>Inouye et al. [53]</td>
<td>Medical and surgical inpatients Over 70 years, n = 776</td>
<td>78, ±6.1 SD</td>
<td>55</td>
<td>Expert opinion using MMSE and BDRS</td>
<td>10</td>
<td>Time and change test Either component incorrect</td>
<td>86</td>
<td>71</td>
</tr>
<tr>
<td>Harwood et al. [54]</td>
<td>Medical inpatients Over 65 years, n = 201</td>
<td>76, range 65–97</td>
<td>51</td>
<td>DSM-III-R</td>
<td>10 dementia alone 13 dementia with delirium</td>
<td>96</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Jitanpunkal et al. [55]</td>
<td>Geriatric inpatients Over 60 years, n = 168</td>
<td>82±6.6 SD</td>
<td>59</td>
<td>DSM-III-R</td>
<td>35</td>
<td>AMTS &lt;8</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td>Erkinjunti et al. [56]</td>
<td>Medical inpatients Over 65 years, n = 282</td>
<td>75, ±7.2 SD</td>
<td>61</td>
<td>Interview and expert opinion</td>
<td>9.6</td>
<td>AMTS &lt;7</td>
<td>81</td>
<td>85</td>
</tr>
<tr>
<td>Klein et al. [57]</td>
<td>Medical inpatients Over 40 years, but results quoted for &gt;60 years</td>
<td>72 demented 66 non-demented</td>
<td>59</td>
<td>DSM-III</td>
<td>52</td>
<td>SPMSQ 3 errors</td>
<td>86</td>
<td>99</td>
</tr>
<tr>
<td>Anthony et al. [58]</td>
<td>Medical inpatients Over 20 years, but results quoted for &gt;60 years n = 97, n = 41 for those &gt;60</td>
<td>No mean age quoted</td>
<td>63</td>
<td>DSM-III</td>
<td>38 dementia and ‘delirium with dementia’</td>
<td>91</td>
<td>85</td>
<td></td>
</tr>
</tbody>
</table>

♀, female; Prev, prevalence of dementia; Sen, sensitivity; Sp, specificity; SD, standard deviation; SE, standard error; DSM, diagnostic and statistical manual of mental disorders; AMTS, Abbreviated Mental Test score; IQCODE, informant questionnaire of cognitive decline in the elderly; MMSE, Mini-Mental State Examination; SPMSQ, short portable mental status questionnaire.
The number of studies we have reported is small, with only the AMTS having more than a single paper investigating its properties. We found only a single study validating the full MMSE in hospital inpatients, despite this being a very common instrument and used as a ‘reference’ standard in six excluded studies. Predictably, as sensitivity increases, the specificity of MMSE reduces. With reference to the instruments recommended in the UK NICE guidelines the MMSE has been validated as described, and we could find no validation for the 6-CIT, GP-COG or 7-minute screen. From the American Geriatrics Society guidance, again only the MMSE has been validated in the hospital population. The British Geriatrics Society guidelines use validated instruments, though the clock-drawing test has only been validated against the MMSE as a reference standard. Other tests in common use where there are no studies validating use in inpatient populations include the MOCA and the Addenbrooke’s cognitive assessment revised version (ACE-R) [32]. We contacted the authors of the MOCA, the Mini-Cog and the ACE-R who confirmed they knew of no validation studies specifically among older inpatients.

The prevalence of dementia in the studies varies from 6 to 59%. This may reflect the changing demographics of acute hospital admissions over time, but must also reflect varying patient selection in the studies included. A carefully conducted and influential recent study [17] reported the prevalence of dementia among unselected older medical admissions at the higher end of this range.

In some of the studies, the sample population was mixed with outpatients and on contacting the authors we were able to extract data from one study [50]. The AMTS was the only instrument which had more than a single study examining its properties. We note that at a cut-off of <7, that across the three hospital inpatient samples reported, very similar sensitivities and specificities were found, although a cut-off of <8 is considered more usual in clinical practice. According to

---

Table 2. Meta-analysis of data from Antonelli Inclaz et al. [52], Harwood et al. [54] and Jitanpunal et al. [55], using the AMTS against reference standard (DSM-IIIR)

<table>
<thead>
<tr>
<th>Test and cut-off</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR− (95% CI)</th>
<th>DOR</th>
<th>I² (%) across all AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMTS &lt;7</td>
<td>0.81 (0.76–0.86)</td>
<td>0.84 (0.83–0.85)</td>
<td>5.04 (4.54–5.61)</td>
<td>0.23 (0.17–0.29)</td>
<td>22.45 (15.92–31.65)</td>
<td>0</td>
</tr>
</tbody>
</table>

LR+, positive likelihood ratio; LR−, negative likelihood ratio; DOR, diagnostic odds ratio; I², test of heterogeneity; AUC, area under the curve from a summary receiver operating characteristic curve.

Figure 2. Summary receiver operating characteristic for three studies using AMTS cut-off of <7 against DSM-IIIR dementia demonstrating an AUC of 0.87. The red circle represents the Q*, where sensitivity equals specificity.

---

Discussion

The number of studies we have reported is small, with only the AMTS having more than a single paper investigating its properties. We found only a single study validating the full MMSE in hospital inpatients, despite this being a very common instrument and used as a ‘reference’ standard in six excluded studies. Predictably, as sensitivity increases, the specificity of MMSE reduces. With reference to the instruments recommended in the UK NICE guidelines the MMSE has been validated as described, and we could find no validation for the 6-CIT, GP-COG or 7-minute screen. From the American Geriatrics Society guidance, again only the MMSE has been validated in the hospital population. The British Geriatrics Society guidelines use validated instruments, though the clock-drawing test has only been validated against the MMSE as a reference standard. Other tests in common use where there are no studies validating use in inpatient populations include the MOCA and the Addenbrooke’s cognitive assessment revised version (ACE-R) [32]. We contacted the authors of the MOCA, the Mini-Cog and the ACE-R who confirmed they knew of no validation studies specifically among older inpatients.

The prevalence of dementia in the studies varies from 6 to 59%. This may reflect the changing demographics of acute hospital admissions over time, but must also reflect varying patient selection in the studies included. A carefully conducted and influential recent study [17] reported the prevalence of dementia among unselected older medical admissions at the higher end of this range.

In some of the studies, the sample population was mixed with outpatients and on contacting the authors we were able to extract data from one study [50]. The AMTS was the only instrument which had more than a single study examining its properties. We note that at a cut-off of <7, that across the three hospital inpatient samples reported, very similar sensitivities and specificities were found, although a cut-off of <8 is considered more usual in clinical practice. According to
our meta-analysis, it demonstrates good properties for a screening instrument, with good sensitivity and specificity (>0.8), a diagnostic odds ratio of >20 and a positive and negative likelihood ratio of >5 and <0.3, respectively. These LR cut-offs are what is considered an accurate test, and the DOR cut-off suggests the instrument is useful in clinical practice. The AUC value of 0.87 is also considered good.

The AMTS has practical limitations. The long-term memory question (When did the first world war end?) is culturally specific and the recognition question requires two people to be at the bed space during the assessment. It does not require pen or paper so is suitable for people with visual or physical impairment. It is also brief, taking 3–4 min. Compared with MMSE, it is shorter, has been validated in more general hospital studies, and shows superior specificity (fewer false negatives). It is also freely available for use without copyright restrictions, as opposed to the MMSE. Used in a general hospital with 1,000 beds, and 10,000 non-elective admissions of people over 65 years per year, assuming all screened and 40% prevalence for dementia, 4,200 (42%) would screen positive, 3,240 (32.4%) would be true positives and 960 (9.6%) false positives. Of the true positives, 1,620 (16.2%) would be already known to services and 1,620 (16.2%) would then require further assessment. Therefore, a total of 2,580 further assessments would be needed. Assuming 60 min each per assessment, either in hospital or after discharge, this would lead to considerable increased resource implications, as well as the false positive patients (960) having potentially stressful assessments.

Our review has limitations. These include publication bias, whereby some studies that may show poor performance of our selected instruments may not have been published. Selection bias can also be a problem in meta-analysis, but we have minimised this by having a strict protocol and solving selection disagreements by consensus. Even using gold standard reference criteria the prevalence of dementia can vary widely depending on the diagnostic criteria used [61], for example, DSM IIIR and DSM IV, and this makes combining the data a challenge. The exclusion of non-English language studies also has potential for bias.

If screening is chosen, timing matters. Inpatient screening needs to take place after the initial acute illness has improved and a diagnosis of delirium has been excluded. As with any brief screen, a second stage procedure is always needed. Such a second stage would involve broadening assessments beyond cognitive tests. Ultimately, screening for dementia always has to be followed by a detailed expert assessment before diagnosis. In many cases such assessment may only be realistic after the hospital admission; screening identifies those warranting such detailed assessment.

At present, evidence that screening for dementia is effective is lacking [9]. If screening is chosen, instruments used need to be short, valid, reliable, acceptable and show good sensitivity and specificity. Given the high prevalence of dementia in hospitals then any screening instrument used must have been validated appropriately in this specific setting, as opposed to using tools validated in community settings where prevalence rates are much lower [62]. Few instruments have been researched in this setting, and further research is clearly needed, but available data should at least allow some predictions about impacts on service of systematic screening. Assuming a prevalence of 40% then any future validation study would need to recruit 154 patients (61 with prevalent dementia) for a study to be powered to 95% confidence ±10% [63]. Any future research should explicitly report timing of screening during an admission, and to minimise bias should ideally ensure rater blindness between the index test and the gold standard test.

### Conclusion

Many instruments are recommended for screening for dementia. A small number have been validated in general hospital inpatients. The AMTS, a fast and commonly used screening test, is currently the most researched instrument for this population, but the review is unable to recommend a single best instrument. There is a clear need for more robust evidence to best inform screening for dementia in hospital inpatients.

### Key points

- Numerous tools are advocated to detect dementia and cognitive impairment.
- Dementia is common and underdiagnosed in the hospital setting, so an admission is a potential opportunity to detect dementia.
- Six discreet instruments have established properties in screening for dementia in hospital inpatients and only one, the AMTS, in more than one report.
- Available data allow prediction of service and patient impacts of any screen for dementia among older inpatients.
- There is a clear need for robust validation studies of dementia screening instruments in hospital inpatients.

### Acknowledgements

We acknowledge the helpful comments on the manuscript of Prof. John Gladman, Professor of the Medicine of Older People, University of Nottingham as well as Ms Nancy Graham of the University of Birmingham Medical School Library for help with devising the search strategy.

### Funding

This work was supported jointly by the Research into Ageing Fund, a fund set up and managed by Age UK and the British Geriatric Society through a grant to TJ. They played no role in the design, execution, analysis and interpretation of data, or writing of the study.
Supplementary data

Supplementary data mentioned in the text is available to subscribers in Age and Ageing online.

References

The very long list of references supporting this review has meant that only the most important are listed here and are represented by bold type throughout the text. The full list of references is available in Supplementary data in Age and Ageing online, Appendix.


Received 8 April 2013; accepted in revised form 26 July 2013