Screening tools to identify hospitalised elderly patients at risk of functional decline: a systematic review

Sutton M, Grimmer-Somers K, Jeffries L

CRD summary
The authors concluded that no gold standard tool could be identified to screen for risk of functional decline in elderly patients admitted to hospital emergency departments. Given the lack of data available, the conclusion is likely to be reliable despite some methodological limitations in the review.

Authors’ objectives
To identify clinically useful screening tools for risk of functional decline in elderly patients presenting to hospital acute care settings.

Searching
MEDLINE, CINAHL, AgeLine, PsychINFO, EMBASE, The Cochrane Library and Cochrane Central Register of Controlled Trials were searched from 1990 to November 2007; the search strategy was reported in full and included a restriction to English-language articles. The bibliographies of retrieved articles were searched for additional studies.

Study selection
All primary study types that reported evidence on the predictive validity, generalisability, reliability or clinical utility of screening tools were eligible for inclusion if they aimed to identify those at risk of functional decline or reduced functional status. Included studies had to be of participants older than 65 years who were admitted to the emergency department of an acute care hospital. Included studies had to report data on the tool’s predictive ability (sensitivity, likelihood ratios, receiver operating characteristic (ROC) or data to calculate these parameters). The five screening tools included in the review incorporated a wide range of items (details reported in the paper); three included Mini Mental State (MMSE) as a measure of cognitive impairment; three were clinician administered and two were clinician administered or self-reported. Maximum total scores ranged from four to 11.5 (high scores indicated increased risk). The reference standards used to verify functional decline also varied and, where reported, were multi-component.

The authors stated neither how studies were selected for the review nor how many reviewers performed the selection.

Assessment of study quality
The methodological quality of included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool, which was used to generate an overall score out of 14 (items designated unclear were scored zero).

The authors did not state how many reviewers performed the quality assessment.

Data extraction
Data were extracted on the items included in the screening tools and their scoring systems and diagnostic thresholds, reference standards used to verify functional decline and measures of predictive validity, reliability, generalisability and clinical utility.

The data extraction tool was developed by one reviewer and validated by a second. The authors did not state how many reviewers performed the data extraction.

Methods of synthesis
The included studies were combined in a narrative synthesis structured by outcome measure (predictive validity, reliability, generalisability and clinical utility). Characteristics of the individual screening tools and measures of their

Database of Abstracts of Reviews of Effects (DARE)
Produced by the Centre for Reviews and Dissemination
Copyright © 2019 University of York
performance were presented in tables.

**Results of the review**

Six studies (n=approximately 4,000, sample size 193 to 1,854) of five screening tools were included in the review. QUADAS scores ranged from 12 to 14 (one study scored 14). Blinding of assessors was generally unreported. No study validated any instrument independently of its development setting.

The end-point for functional status determination by reference standard varied across the studies (one to six months, where reported). Area under the ROC curve was reported for four of the five screening tools (four studies, one study per tool). Area under curve was 0.64 and 0.66 for the TRST tool (30 day and 120 day end-point), 0.65 for the HARP tool (discharge end-point), 0.71 for the ISAR tool (six month end-point) and 0.73 for the SHERPA tool (three month end-point). One study reported only sensitivity (88%) and specificity (54%) for a low cut-off, graded 0 (low risk) out of 4.

One study reported the reliability of the ISAR instrument: concordance correlation coefficient for test-retest reliability was 0.73.

The paper also included a discussion of generalisability and clinical utility, but no supporting data were available.

**Authors’ conclusions**

No single tool reported sufficient predictive validity to warrant recommendation over the others for screening for risk of functional decline in elderly patients admitted to hospital acutely.

**CRD commentary**

The review stated a clear research question and defined appropriate inclusion criteria. The search strategy was fully reported and wide ranging, but the restriction to English-language publications may have resulted in the omission of relevant studies and left the review open to language bias. The review process was not fully reported and the potential for error and/or bias could not, therefore, be assessed. The methodological quality of included studies was assessed, but the QUADAS tool was used inappropriately to generate overall quality scores, which limited the informative value of the quality assessment. Importantly, no study validated any tool in a population independent from that used to develop it and so the estimates of predictive value presented may be optimistic. The use of a narrative summary was appropriate given the paucity of data and wide variation in study characteristics. Given the lack of evidence available, the authors’ conclusion that no gold standard screening tool could be identified was reasonable.

**Implications of the review for practice and research**

**Practice:** The authors stated that the potential for significant numbers of false positive and negative results and lack of validation in different hospital settings meant that the use of interventions based solely on assessments using any of these tools would be questionable.

**Research:** The authors stated that further research was needed to determine the accuracy and reliability of screening tools in different healthcare settings, and to assess the potential of adjunctive procedures to improve their predictive value.

**Funding**

Central Northern Adelaide Health Service.

**Bibliographic details**


**PubMedID**

19166437
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.