Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults

Gates S, Smith LA, Fisher JD, Lamb SE

CRD summary
This review summarised evidence on accuracy of screening tools for predicting risk of falling in community-living older adults. The authors concluded that insufficient evidence existed to show that any screening instrument was adequate for predicting falls. Despite some limitations in the reporting of the review, these conclusions reflect the evidence presented.

Authors’ objectives
To summarise the evidence on the accuracy of screening tools for predicting risk of falling in community-living older adults.

Searching
MEDLINE, EMBASE, PsycINFO, CINAHL and Science Citation Index were searched from inception to May 2007 for relevant studies. Search terms were reported. Reference lists of six previously published reviews and of retrieved studies were examined for further relevant evidence.

Study selection
Prospective cohort studies that evaluated one or more screening tests for predicting falls among elderly people living in the community or substantially independently were eligible for inclusion in the review. Studies had to record falls prospectively (with a follow-up duration of at least three months) and report sufficient data to allow calculation of diagnostic accuracy measures.

Mean age of participants among included studies ranged from 71.5 to 84.7 years, where reported. Where reported, between 10.6% and 50.6% of participants had a previous fall. Twenty-nine different screening tests were evaluated, most frequently Tinetti gait, balance and mobility scales, and Timed Up and Go (TUG) test. Other tests included Mobility Interaction Fall Chart, functional reach, tandem stance and walking tests. Length of follow-up ranged from three months to five years.

The authors did not state how many reviewers selected the studies for inclusion.

Assessment of study quality
Studies were evaluated using a modified version of the Quality Assessment of Diagnostic Accuracy Studies (QUADA) criteria. This modified version contained 16 criteria related to the quality of study conduct, each of which could be met fully, partially or not.

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

Data extraction
Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) measures and their related 95% confidence intervals (CIs) were either extracted directly from included studies or calculated from reported 2x2 data. Where studies reported area under the curve (AUC) for summary receiver operating characteristic (ROC) curves, these data were extracted.

Two reviewers independently extracted data from included studies. Disagreements were resolved by consensus or through consultation with a third reviewer.
Methods of synthesis
Due to clinical and methodological heterogeneity, studies were combined in a narrative synthesis, grouped by type of test evaluated.

Results of the review
A total of 25 diagnostic cohort studies (n>5,920) were included in the review. Methodological quality was variable. For example, most studies (21 out of 25) reported details of screening test procedures, but few (eight out of 25) reported results for all the screening tests performed. Ten of the 25 studies had errors in their reported analyses. Loss to follow-up ranged from 0% to 33.5%.

Eight studies evaluated the Tinetti gait, balance and mobility scales. Sensitivity among five studies that evaluated the overall mobility scale ranged from 0.27 to 0.76 and specificity ranged from 0.52 to 0.83. Seven studies that evaluated the balance scale alone reported sensitivity values of 0.08 to 0.80 and specificity values of 0.66 to 0.97; one study reported an AUC of 0.56. Four studies evaluating the gait scale alone and reported sensitivity values of 0.20 to 0.68 and specificity values of 0.63 to 0.95.

Three studies of the TUG test reported sensitivity values of 0.10 to 0.54 and specificity values of 0.73 to 0.95; one study reported an AUC of 0.61. Tests and cut-offs varied between these studies.

Results of studies that evaluated Mobility Interaction Fall Chart, functional reach, tandem stance and walking tests were inconsistent (two studies each).

All other tests were evaluated in single studies.

Authors’ conclusions
Insufficient evidence existed that any screening instrument was adequate for predicting falls.

CRD commentary
This review was based on a question clearly defined in terms of the participants, interventions, outcomes and study designs of interest. The search covered multiple electronic databases and other sources. It was unclear whether there were any language restrictions on this search and so the potential for language bias could not be entirely ruled out. Attempts were made to minimise errors and bias in data extraction; it was unclear whether similar efforts were made in study selection or validity assessment. Validity was assessed using an appropriate method. Use of a narrative synthesis was appropriate given heterogeneity between identified studies. Despite the limitations mentioned above, the authors’ conclusions follow from the evidence presented. Interpretation of the conclusions should take into account limitations in the included studies and the reported review process.

Implications of the review for practice and research
The authors did not state any implications for practice.

Research: The authors stated that further studies were required. These should have a sufficiently large sample size to estimate sensitivity and specificity with precision, be conducted in a clinically relevant population with a sufficient duration of follow-up and use reliable methods for recording falls. The authors added that individual patient data from existing risk factor studies may also be necessary.

Funding
UK National Health Service, Service Delivery and Organisation Programme, project number SDO/139/2006.

Bibliographic details