Published in final edited form as:
Age Ageing. 2008 November ; 37(6): 621–627. doi:10.1093/ageing/afn203.

A systematic review and meta-analysis of studies using the STRATIFY tool for prediction of falls in hospital patients: how well does it work?

David Oliver[1], **Alexandra Papaioannou**[2], **Lora Giangregorio**[3], **Lehana Thabane**[4], **Katerina Reizgys**[5], and **Gary Foster**[6]

[1]School of Health and Social Care (& Institute of Health Sciences), University of Reading, London Road, Reading, RG1 5AQ, UK

[2]Department of Medicine, McMaster University, Hamilton, Ontario, Canada

[3]Department of Kinesiology, University of Waterloo, Canada

[4]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada

[5]Schulich School of Medicine and Dentistry, University of Western Ontario, Canada

[6]Centre for Evaluation of Medicines, St Joseph's Healthcare, Hamilton, Ontario, Canada

Abstract

Background—STRATIFY is a prediction tool developed for use in for hospital inpatients, using a 0–5 score to predict patients who will fall. It has been widely used as part of hospital fall prevention plans, but it is not clear how good its operational utility is in a variety of settings.

Objectives—(i) to describe the predictive validity of STRATIFY for identifying hospital inpatients who will fall via systematic review and descriptive analysis, based on its use in several prospective cohort studies of hospital inpatients; (ii) to describe the predictive validity of STRATIFY among inpatients in geriatric rehabilitation via meta-analysis and (iii) in turn, to help practitioners and institutions wishing to implement interventions to prevent in-hospital falls.

Methods—a systematic literature review of prospective validation studies of STRATIFY for falls prediction in hospital inpatients. For inclusion, studies must report prospective validation cohorts, with sufficient data for calculation of sensitivity (SENS), specificity (SPEC), negative and positive predictive value (NPV and PPV), total predictive accuracy (TPA) and 95% confidence intervals (CI). We performed meta-analysis using precision-weighted fixed- and random-effects models using studies that evaluated STRATIFY among geriatric rehabilitation inpatients.

Address correspondence to: Dr D. Oliver. Tel: (+44) 1183786851; Fax: (+44) 1183786862. d.oliver@reading.ac.uk.

Supplementary data
Supplementary data for this article are available online at http://ageing.oxfordjournals.org.

Conflicts of interest
This study received no external research funding or commercial sponsorship. The authors have no conflicts of interest to declare.

CIHR Author Manuscript

CIHR Author Manuscript

Measurements—key features of the patient population, setting, study design and numbers of falls/fallers were abstracted. SENS, SPEC, PPV, NPV, TPA and 95% CI were reported for each cohort. Pooled values and chi-squared test for homogeneity were reported for a meta-analysis of studies conducted in geriatric rehabilitation settings.

Results—forty-one papers were identified by the search, with eight ultimately eligible for inclusion in the systematic review and four for inclusion in the meta-analysis. The predictive validity of STRATIFY, using a random-effects model, for the four studies involving geriatric patients was as follows: SENS 67.2 (95% CI 60.8, 73.6), SPEC 51.2 (95% CI 43.0, 59.3), PPV 23.1 (95% CI 14.9, 31.2), NPV 86.5 (95% CI 78.4, 94.6). The $Q_{(3)}$ test for homogeneity was not significant for SENS at $P = 0.36$, but it was significant at $P < 0.01$ for SPEC, PPV and NPV. TPA across all four studies varied from 43.2 to 60.0.

Conclusion—the current study reveals a relatively high NPV and low PPV and TPA for the STRATIFY instrument, suggesting that it may not be optimal for identifying high-risk individuals for fall prevention. Further, the study demonstrates that population and setting affect STRATIFY performance.

Keywords

falls; hospital; predictive validity; STRATIFY; systematic review; elderly

Introduction

Falls the commonest safety incident among hospitalised patients. They account for 32% of incident reports in UK hospitals [1]. Rates from 5 to 18 falls per 1000 bed days have been described in intervention [2] and observational studies. [3, 4, 5, 6, 7, 8]—all a probable underestimation of the true incidence [1, 9–12]. Falls in hospital may lead to injury in up to 30% of cases and associated mortality and morbidity [1, 4, 7, 13]. They may cause anxiety, [14–16], loss of confidence, impaired rehabilitation and function, prolonged stay and discharge to long-term care settings. [17–19]. Falls also cause guilt and anxiety for staff [1–4], complaint or litigation from patients' relatives who may feel that no falls are acceptable. [4, 11, 20, 21]. Falls in hospital are associated with excess financial and opportunity costs [17, 19].

A recent major systematic review of interventions to prevent falls in hospital [2] identified some benefit from multifaceted interventions in hospital settings on falls rates (though not rates of injury or individual risk of falling). The components of these interventions and the populations studied vary, and not all positive studies have employed formal falls risk prediction tools such as STRATIFY, instead of focusing on post-fall assessment or identification and intervention for falls risk factors. Despite the lack of evidence for falls risk assessment tools, many hospitals continue to employ them [1, 4, 6]. However intuitively attractive the use of such tools might be, if they do not perform sufficiently well in that setting and population, their use may be ineffectual or provide false reassurance that 'something is being done' to target high-risk patients whilst diverting attention away from more potentially effective interventions [22, 23].

We must distinguish between risk tools which identify risk factors to prompt action for each; those which produce a continuous score which can be used to estimate cumulative risk and those which yield categorical 'yes/no' predictions of 'high' or 'low risk'—it is this third type of risk tool which this study examines. To be operationally useful, a prediction tool should ideally have the characteristics [23, 24], initially set out by Wyatt and Altman [25] and re-emphasised in the 'Standards for Reporting of Diagnostic Accuracy' (STARD) guidance on clinical assessment and diagnostic tools [26], and re-iterated in a recent systematic review [27]: ease of completion; high adherence by staff; high inter-rater reliability; a transparent calculation of risk score based on the operational properties of the tool (and not arbitrarily assigned values); face validity; prospective validation for predictive validity [sensitivity (SENS), specificity (SPEC), positive predictive value (PPV), negative predictive value (NPV) and total predictive accuracy (TPA)]—ideally with sufficient power to allow narrow confidence intervals for these values; validated in more than one cohort of patients—and preferably in a cohort and setting similar to that for which their operational use is intended (external validity). The tool should ideally confer greater accuracy than the 'best guess' professional judgment of staff on the wards if it is to add value to clinical assessment. But how many hospital falls risk assessment tools for use in hospital even begin to meet these criteria?

Four systematic reviews [24, 27–29] of hospital falls prediction tools have identified only two (STRATIFY [30] and Morse Falls Scale [11]) which have been subjected to prospective validation in two or more cohorts, with appropriate tests of predictive validity, though in a recent 'head to head' comparison within the same patient cohort, STRATIFY [31] proved to be more accurate and more frequently completed. STRATIFY was originally derived in mixed acute/rehabilitation geriatric wards of a UK urban teaching hospital using a 'case-control' design and multivariate regression to identify predictors of falls in hospital inpatients. This resulted in a simple five-point score (each item scoring 1 or 0), with predictive "cut-offs" as 2 or 3 used in the original validation studies which followed. STRATIFY was not designed or validated for continuous modelling of risk but for use in categorical prediction 'high' versus 'low risk'. Sensitivity and specificity were both found to be in excess of 80% in the two UK cohorts of the original paper, leading to widespread adoption of the tool in clinical practice. It is now 10 years since the publication of the original STRATIFY paper and a number of prospective studies in several cohorts of patients have been published.

We report a systematic review and meta-analysis of the operational utility of the STRATIFY fall prediction tool. The aims of the study were (i) to describe the predictive validity of STRATIFY for identifying hospital inpatients who will fall via systematic review and descriptive analysis, based on its use in several prospective cohort studies of hospital inpatients, and (ii) to describe the predictive validity of STRATIFY among inpatients in geriatric rehabilitation via meta-analysis. This should in turn help practitioners and institutions wishing to implement interventions to prevent inhospital falls.

Methods

Study search and selection

In line with QUOROM guidelines [32] on the conduct of systematic reviews, a literature search was conducted for articles published from 1997 (the year of initial publication of STRATIFY) to February 2006: Ovid MEDLINE, EMBASE, AARP Ageline, CINAHL, CDSR, ACP Journal Club, DARE and CCTR. The search strategy included the terms 'STRATIFY' and either 'falls', 'risk assessment', 'clinical assessment tools' or 'inpatients' as the keywords. Additional articles were identified by hand searching of bibliographic references, including reference to published systematic reviews of falls risk assessment tools for hospital patients [24, 27–29] and prevention strategies for falls in hospital [2, 4, 8]. Abstracts of all articles identified by the search were reviewed; for those that fulfilled the eligibility criteria the full text articles were retrieved.

Inclusion criteria and quality assessment

All primary studies including a prospective validation of STRATIFY for predicting falls in hospital inpatients were retrieved. To be included, studies must have been published in a peer-reviewed journal or letter to the editor containing original data. Using the number of 'fallers' or 'falls', each study must also have included sufficient data to calculate the sensitivity, specificity, positive and negative predictive values, odds ratio and confidence intervals. Among those, only papers where the population was exclusively geriatric rehabilitation patients, where the predictive validity was for patients who fell (rather than falls as discrete events) and which used a STRATIFY score cut-off set at greater than or equal to two were included in the pooling for meta-analysis.

Articles were excluded if they were published in non-peer reviewed journals, or included non-original data, or if not exclusively describing hospital inpatients.

Abstraction of data and analysis

Two assessors independently abstracted data from each study. Patient demographic data, any relevant clinical characteristics and details of the study design were described. Outcomes of interest included number of falls, number of fallers (patients who fell at least once during validation), as well as the following data specific to the predictive ability of the STRATIFY instrument: SENS, SPEC, PPV, NPV, odds ratio (OR) and 95% confidence intervals (CI). Any discrepancies were resolved by consensus.

Statistical analyses

This study was conducted according to the Cochrane Methods Working Group guidelines for review and meta-analysis of diagnostic studies [33]. The retrieval process was illustrated using a flow diagram in line with QUOROM guidance on the reporting of systematic reviews [32] (Figure 1). The eight studies listed in Table 1 met the inclusion criteria for the systematic review [30, 31, 34–39]. Milisen *et al.* [39] reported on two separate cohorts, one surgical and one geriatric. The results of the geriatric cohort were included in the meta-analysis but those for the surgical cohort were not.

The original local and remote validation cohorts for STRATIFY reported by Oliver *et al.* [30], whilst both describing large cohorts of patients and high values for SENS, SPEC and NPV, described the predictive validity of STRATIFY for 'falls' as discrete events rather than 'fallers' or 'patients who fell' (in contrast to all other papers identified) and so could not be pooled with other studies for meta-analysis.

Estimates with corresponding 95% CI for SENS, SPEC, PPV and NPV were computed for each study separately (see Table 2) [40]. SENS (or the 'true positive rate') is the percentage of patients who fell and had been identified as 'high risk'. SPEC (or the 'true negative rate') is the percentage of patients who did not fall and had been identified as 'low risk'. PPV is the percentage of patients identified as 'high risk' who went on to fall. NPV is the percentage of patients identified as 'low risk' who went on not to fall [24, 41]. Estimates of SENS, SPEC, PPV and NPV were pooled or combined using the precision-weighted fixed- and random-effects models [42, 43]. A chi-squared (i.e. Q-test) test was used to assess between-study heterogeneity for each outcome measure using $\alpha = 0.10$ as the criterion for statistical significance [44]. The results are reported as the pooled estimates, corresponding 95% confidence interval and associated P-value for the test of significance. All analyses were performed using SAS 9.1 (Cary, NC).

Results

Studies identified

The initial literature search identified 24 references which appeared to fulfil the inclusion criteria and were requested in full. After further scrutiny, 15 were rejected and 10 included for further analysis [30, 31, 34–39, 45, 46]. One article by Smith *et al.* [45] was subsequently excluded from the systematic review because patients with a length of stay >28 days were excluded from its study population, and therefore the sample may not have been representative of the target sample. The study of Jester *et al.* [46] was excluded from the review because it described only retrospective fitting of STRATIFY to a small sample of patients admitted with hip fracture.

The eight studies included in the systematic review are set out in Table 1. The references not included, along with the reasons for rejection, are listed as supplementary references [S1–S30], available at *Age and Ageing* online. The search also identified three systematic reviews on falls risk assessment [24, 28, 29] already cited in the Introduction.

The key features of these eight studies are set out in Table 1. Several papers reported more than one validation cohort within the reported study [30, 38, 39]. In addition, several included articles [30, 34–36, 37, 39] reported the predictive ability of STRATIFY for more than one 'cut-off' score.

Although the predictive validity data for all these included studies have been described, not all were eligible for inclusion in final meta-analysis. From meta-analysis, we excluded Papaiaoannou [35], Vasallo [31] and Chiari [38], as they included clinically heterogeneous populations including younger adult acute and surgical patients. In addition, the two validation cohorts reported in [30] in the original STRATIFY paper were also excluded from meta-analysis as they used ‘falls’ rather than patients who fell (‘fallers’) as the basis for predictive validity, in contrast to all identified studies. Table 3 summarises predictive validity analyses for these four papers. The meta-analysis was conducted with data where a cut-off score of 2 was used. The methods used for dealing with statistical heterogeneity were contingent on the results and so are reported below. All four of the studies included in the meta-analysis (Coker [34], Haines [36], Hill [37] and Milisen [39]) incorporated geriatric rehabilitation patients exclusively. Table 2 summarises the predictive validity for each cohort and the meta-analyses. The populations in [35], [31] and [38] were heterogeneous, and Papaioannou, Coker, Hill and Haines all reported results using more than one STRATIFY score cut-off.

Predictive validity for hospital inpatient fallers

The key operational properties for each prospective cohort study are summarised in Table 2. For predicting fallers, SENS varied from 42.9% [37] to 76.9% [36]. SPEC varied from 43.2% [37] to 59.0% [39]. PPV varied from 12.5% [37] to 29.9% [34, 36]. NPV varied from 80.0% [37] to 93.0% [39]. The original STRATIFY validation cohorts [30] report considerably better predictive ability, but in these cohorts completion of the STRATIFY score was performed weekly, and analysis of predictive validity was performed for falls in that week rather than for fallers. TPA varied from 43.2% [37] to 60.0% [39]. The operational properties varied widely between populations and cut-off scores employed.

Meta-analysis of predictive validity for fallers in inpatient geriatric rehabilitation

The four studies pooled for meta-analysis [34, 36, 37, 39] were clinically homogeneous in terms of age and settings. However, the $Q_{(\mathrm{df}=3)}$ test for heterogeneity was significant for SPEC, PPV and NPV ($P < 0.01$), but it was not significant for SENS ($P = 0.36$). Because the test for heterogeneity was statistically significant for three of the four estimates, we conclude that sufficient between-study heterogeneity was present to warrant using a random-effects model. Based on the random-effects model using STRATIFY to predict inpatient falls, pooled estimates (95% CI)were 67.2 (60.8, 73.6) for SENS, 51.2 (43.0, 59.3) for SPEC, 23.1 (14.9, 31.2) for PPV and 86.5 (78.4, 94.6) for NPV.

Discussion

The idea of using a falls prediction tool to target inpatients for fall prevention strategies is an attractive one to organisations and clinicians. Tools have frequently been used both in research and ‘real life’ intervention programs, with the STRATIFY tool being widely adopted. The current study summarised the predictive validity of STRATIFY in predicting patients who would fall (rather than falls as discrete events) in eight published studies using variously scores of 2 or 3 as the ‘cut-off’ for prediction and also in one published study [30] describing two validation cohorts where ‘falls’ rather than ‘fallers’ were the event predicted. The operational properties of the tool varied considerably according to population and setting. In particular, PPV and NPV were dependent on the prevalence of the index condition (in this case falls) in the population. Furthermore, elements of the STRATIFY score (e.g. agitation, gait instability, previous fall) might vary in prevalence between different inpatient populations (e.g. acute, rehabilitation or mental health). Four of the included studies were pooled for meta-analysis though the pooled specificity and positive predictive value in particular were only modest, making the tool of doubtful utility in accurately identifying potential fallers for preventative interventions.

The initial publication in 1997 of the STRATIFY score by Oliver *et al.* [30] showed very high SENS, SPEC and NPV for falls as discrete events in two UK hospital populations of mixed acute/rehabilitation geriatric patients, though much lower PPV. These findings were strengthened by a subsequent systematic review [24] comparing it favourably to other tools and in a head-to-head comparison with another tool in one cohort of patients [31]. STRATIFY showed a higher ease of completion, inter-rater reliability and predictive validity. Collectively, these studies and the simplicity of STRATIFY implementation (five ‘yes’ or ‘no’ items with unweighted 0–5 scoring and no need for specialised clinical assessment) have influenced its adoption in the clinical realm. However, the original 1997 study [30] reported the validation of STRATIFY in predicting falls rather than fallers and was in settings and populations similar to those in which the tool had initially been developed. Such tools always run the risk of being less powerful when translated to other settings, reducing their external validity.

This study reports the first published meta-analysis of a hospital falls risk prediction tool in the geriatric rehabilitation population. The search strategies, inclusion criteria and abstraction of data adhered to recognised methodological standards. There were inherent methodological problems. Firstly, despite the large number of studies initially identified, only a small number actually reported the data in a way which made inclusion for further analysis possible. Secondly, we were forced to exclude the paper [30] reporting the initial development and validation cohorts for the reasons outlined. Thirdly, we analysed published aggregated data rather than individual patient data.

The findings of the current study have important implications for clinical practice. First, any falls risk prediction tool should be validated in the setting and population (or similar) in which it is to be used. Secondly, even where high NPV or SPEC values may lead to reasonable reassurance that some patients are at low risk of falls, if PPV is low, then there is a risk that falls interventions will be poorly targeted with most patients being deemed ‘high risk’. Third, some components of STRATIFY (e.g. agitation, gait instability, urinary frequency) are not static but may vary from day to day, making a ‘one off’ risk assessment of dubious value. Fourth, we know that in elderly inpatient populations, ~50% of falls occur in patients who have already fallen once, so that this ‘index’ fall should trigger assessment and intervention [1, 3]. Perhaps it would be better to assess all admitted patients for reversible falls risk factors and make a treatment and risk minimisation plan for each of these, and then use each subsequent fall as a trigger for reassessment of risk. Staff and organisations should be made aware of the limitations of risk assessment tools and not seduced by the attractiveness of an ‘off the shelf’ solution to the problem of falls. Indeed, if we examine the higher quality and most successful falls intervention trials in the literature, they have not tended to rely on formal risk prediction but rather use post-fall assessment and identification and treatment/risk minimisation plans for common risk factors in every patient [47–50].

The idea of a falls risk prediction tool which can be adapted and validated for each patient population remains of interest, as does the comparison of falls prevention strategies with and without an element of formal risk assessment, or at least a more rigorous analysis of the relative benefit from various components in the ‘black box’ of multifaceted falls interventions (which may include the use of falls risk assessment tools). However, the operational limitations of such tools in daily practice need to be fully considered. Future research should seek to minimise operational limitations when validating or implementing a falls prediction tool and should determine the predictive validity of the tool in the setting and population of interest, including more than one cohort of patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

(Due to the large number of references, only 30 are listed below and are represented by bold type throughout the text. The full list can be found in the supplementary data at *Age and Ageing* online.)

1. National Patient Safety Agency. Slips Trips and Falls in Hospital. London: NPSA; 2007. Available at www.npsa.nhs.uk [13 July 2007, date last accessed]
2. Oliver D, Connelly J, Victor C, et al. Strategies to prevent falls and fractures in hospitals and care homes and effect of cognitive impairment. Systematic review and meta-analyses. BMJ. 2007; 334:82–7. [PubMed: 17158580]
4. Australian Council on Safety. Best Practice Guidelines. Brisbane: Australian Council on Safety; 2005. Falls prevention in hospitals and residential aged care facilities.
6. Lord, S.Sherrington, C.Menz, H., Close, J., editors. Falls in Older People. 2. Vol. Chapter 18. Cambridge: Cambridge University Press; 2007. Preventing falls in institutions.
8. Registered nurses association of Ontario, Hamilton Canada. Nursing Best Practice Guidelines. Hamilton, CA: RNAO; Prevention of falls and fall injuries in the older adult. (revised March 2005). Available at www.rnao.org/bestpractices/PDF/BPGfalls-rev05.pdf [16 September 2008, date last accessed]
22. Oliver D. Risk assessment tools for falls in hospital inpatients. Time to put them to bed? Age Ageing. 2008; 37:248–50. [PubMed: 18456789]
23. Mayer, D. Essential Evidence-based Medicine. Vol. Chapter 13. Cambridge: Cambridge University Press; 2004. Risk assessment.
24. Oliver D, McMurdo M, Daly F, Martin F. Risk factors and risk assessment tools for falls in hospital inpatients: a systematic review. Age Ageing. 2004; 33:122–30. [PubMed: 14960426]
25. Wyatt J, Altman D. Prognostic models: clinically useful or quickly forgotten? BMJ. 1995; 311:539–41.
26. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. AJR Am J Roentgenol. 2003; 181:51–7. [PubMed: 12818829]
27. Scott V, Votova K, Scanlan A, Close J. Multifactorial and functional mobility assessment tools for falls risk among older adults in the community, home-support, long-term care and acute settings. Age Ageing. 2007; 36:130–40. [PubMed: 17293604]
28. Myers H. Hospital falls risk assessment tools: a critique of the literature. Int J Nurs Pract. 2003; 9:233–5.
29. Perrell KL, Nelson A, Goldman R, et al. Fall risk assessment measures: an analytic review. J Gerontol A Biol Sci Med Sci. 2001; 56:761–6.

30. Oliver D, Britton M, Seed P, Martin FC, Hopper A. Development and evaluation of an evidence-based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. BMJ. 1997; 315:1049–53. [PubMed: 9366729]
31. Vassallo M, Stockdale R, Sharma JC, Briggs R, Allen S. Author: a comparative study of the use of four fall risk assessment tools on acute medical wards. J Am Ger Soc. 2005; 53:1034–8.
32. Moher D, Cook D, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. Lancet. 1999; 354:1896–900. [PubMed: 10584742]
33. Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests. Recommended methods. Available at http://www.cochrane.org/docs/sadtdoc1.htm
34. Coker E, Oliver D. Evaluation of the STRATIFY falls prediction tool on a geriatric unit. Outcomes Manag. 2003; 7:8–16. [PubMed: 12593120]
35. Papaioannou A, Parkinson W, Cook R, Ferko N, Coker E, Adachi JD. Prediction of falls using a risk assessment tool in the acute care setting. BMC Med. 2004; 2:1. [PubMed: 14736342]
36. Haines T. All not lost on falls risk assessment for hip fracture patients. J Orthop Nurs. 2006; 10:57.
37. Hill K, Vrantsidis F, Jessup R, McGann A, Pearce J, Collins T. Design-related bias in hospital falls risk screening tools predictive accuracy evaluations: Systematic review and meta-analysis. Australas J Pordiatr Med. 2004; 38:99–108.
38. Chiari P, Mosci D, Fontana S. Comparison of two scales for the assessment of the risk of falls in hospitalised patients [Italian]. Assist Inferma Ricerca. 2002; 21:117–24.
39. Milisen K, Staelens N, Schwendimann R, De Paepe L, Verhaege J, Braes T, et al. Fall prediction in inpatients using the STRATIFY Instrument: A multi-center study. J Am Ger Soc. 2007; 55(5):725–33.
40. Hays, WL., editor. Statistics for the Social Sciences. 2. Vol. Chapters 12 and 13. New York: Holt, Rinehart, & Winston; 1973.
41. Bowers, D.House, A., Owens, D., editors. Understanding Clinical Papers. Vol. Chapter 14. Chichester: Wiley; 2003. Measuring the characteristics of measures.
42. Midgett AS, Stukel TA, Littenberg B. A meta-analytic method for summarizing diagnostic test performances: receiver-operating characteristics summary point estimates. Med Dec Making. 1993; 13:253–7.
43. Irwig L, Macskill P, Galsziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol. 1995; 48:119–30. [PubMed: 7853038]
45. Smith J, Forster A, Young J. Use of the ‘STRATIFY’ falls risk assessment in patients recovering from acute stroke. Age Ageing. 2006; 35:138–43. [PubMed: 16368736]
46. Jester R, Wade S, Henderson K. A pilot investigation of the efficacy of falls risk assessment tools and prevention strategies in and elderly hip fracture population. J Orthop Nurs. 2005; 9:27–34.
47. Fonda D, Cook J, Sandler V, Baily M. Reducing serious falls-related injuries in hospital. Med J Aust. 2006; 184:379–82. [PubMed: 16618235]

Key points

- Falls are the commonest safety incident in hospital.
- Prevention strategies often employ prediction tools.
- STRATIFY (a five-point score) has been subjected to the most independent validation studies and compares well with other tools on speed, adherence and reliability.
- A systematic review identified nine high-quality independent validation cohorts.
- Although high values were reported for specificity and negative predictive value, sensitivity and positive predictive value were generally too low to make the use of such a tool (or similar ones) operationally useful in falls prevention in hospital.

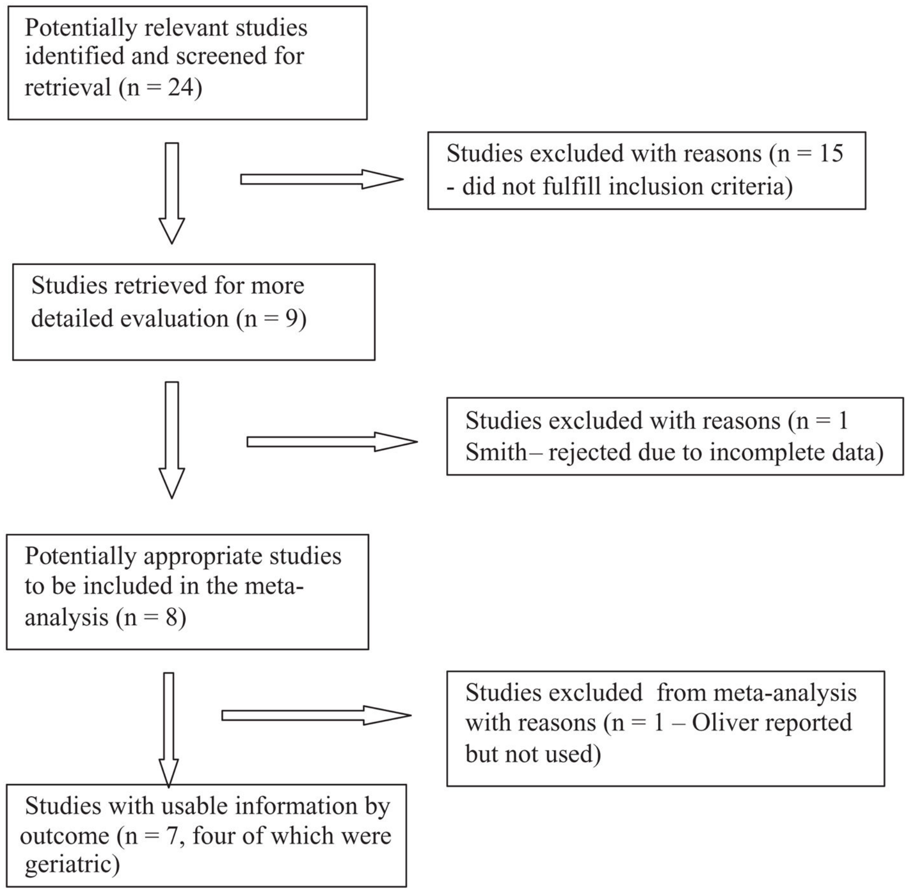

Figure 1.
Flow chart for considered and included studies

Table 1

Trial characteristics for studies included in the systematic review

Trial (first author, year, country)	*n* included	Mean age (years or range)	Time frame patients followed	Population type	Exclusion criteria	Number of falls	Number of fallers
Vassallo, 2005, UK [31][a]	135	83.8	Average LOS 14.6 (±7.5) days	Acute medical wards	Not given	29	22
Coker, 2003, Canada [34]	432	81	Average LOS 50 days	Geriatric rehabilitation	Not given	–	111
Papaioannou, 2004, Canada [35][a]	620	78	Not given	Acute medical wards	Palliative, critical care	77	34
Haines, 2006, Australia [37]	122	79	4004 patient-days	Geriatric rehabilitation	Not given	59	26
Hill, 2004, Australia [37]	44	79.8	Not given	Geriatric rehabilitation	Unable to speak English or Italian, <65 years	–	7
Chiari, 2002, Italy [38][a]	1,181	65–104	2 months	Mixed (internal medicine, geriatric rehab, oncology)	<65 years, hospitalised <48 h	–	51
Milisen, 2007 Belgium [39]	687	67.2	Average LOS 10.2 (±11.4) days	Geriatric and acute l medical wards	Incomplete data, hospitalised <48 h	120	82
Oliver[b], 1997, UK [30]	217	79.5	8 weeks	Geriatric, acutely ill, stroke rehab, disabled	Not given	71	324
Oliver[b], 1997, UK [30]	446	83	8 weeks	Geriatric, acutely ill, stroke rehab, disabled	Not given	79	–

[a]Excluded from final meta-analysis as clinically heterogeneous population, including younger more acute patients;

[b]excluded from final meta-analysis as falls rather than fallers were event used as basis of predictive validity analysis;

LOS = length of stay.

Table 2

Summary of predictive validity analyses for studies included in the systematic review but excluded from pooling for meta-analysis, as not geriatric rehabilitation settings

Study[a]	**SENS(95% CI)**	**SPEC(95% CI)**	**PPV(95% CI)**	**NPV(95% CI)**
Papaioannou [35]	91 (77, 97)	49 (45, 53)	9.4 (7, 13)	99 (97, 100)
Vassallo [31]	68 (47, 84)	66 (57, 74)	28 (18, 42)	91 (83, 96)
Chiari [38]	19 (11, 32)	88 (85, 89)	6 (4, 12)	96 (95, 97)
Oliver [30] (Cohort 1)[b]	93.0 (84., 98)	88 (36, 91.0)	62 (52, 71)	98 (96, 99)
Oliver [30] (Cohort 2)[b]	55 (43, 66)	88 (84, 90)	48 (38, 60)	90 (86, 93)

[a]Whereas more than one cut-off score is reported in the original paper, cut-off with best predictive accuracy is reported in this table;

[b]Oliver reported predictive validity analysis for falls as individual events rather than patients who fell (fallers).

CIHR Author Manuscript
CIHR Author Manuscript

Table 3

Predictive validity analyses for trials included in meta-analysis, fixed- and random-effects meta-analyses and test of homogeneity (patients in geriatric rehabilitation settings)

	A	B	C	D	SENS	SPEC	PPV	NPV
Coker [34]	73	171	38	150	65.8 (56.5, 73.9)	46.7 (41.3, 52.2)	29.9 (24.5, 35.9)	79.8 (73.5, 84.9)
Haines [37]	20	47	6	49	76.9 (58.0, 89.0)	51.0 (41.2, 60.8)	29.9 (20.2, 41.7)	89.1 (78.2, 94.9)
Hill [37]	3	21	4	16	42.9 (15.8, 75.0)	43.2 (28.7, 59.1)	12.5 (4.3, 31.0)	80.0 (58.4, 91.9)
Milisen [39]	55	248	27	357	67.1 (56.3, 76.3)	59.0 (55.0, 62.9)	18.2 (14.2, 22.9)	93.0 (90.0, 95.1)
Fixed-effects estimate [95% CI]					67.2 [61.1, 73.3]	54.1 [51.1, 57.1]	22.5 [19.3, 25.7]	90.5 [88.3, 92.7]
Random-effects estimate [95% CI]					67.2 [60.8, 73.6]	51.2 [43.0, 59.3]	23.1 [14.9, 31.2]	86.5 [78.4, 94.6]
Q (df)[a], *P*-value					3.18 (3), 0.3646	15.17 (3), 0.0017	14.18 (3), 0.0027	18.45 (3), 0.0004
	Fallers	Non-fallers						
STRATIFY cut-off ≥2	A[b]	B[b]						
STRATIFY cut-off <2	C[b]	D[b]						

[a] *Q* = test of homogeneity chi-squared statistic; df = degrees of freedom;

[b] SENS = A/(A + C), SPEC = D/(B + D), PPV = A/(A + B), NPV = D/(C + D).