Title: **The STRATIFY tool and clinical judgment were poor predictors of falling in an acute hospital.**

Authors: Adjunct Professor Joan Webster, RN, BA^{1,2,3}

Professor Mary Courtney, RN, B Commerce, M Health Planning PhD²

Nicole Marsh, RN¹,

Catherine Gale, RN¹

Belynda Abbott RN¹

Anita Mackenzie-Ross B Occ Health¹

Prue McRae. B Physiotherapy¹

¹Royal Brisbane and Women's Hospital

²Queensland University of Technology

³Griffith University

Mailing address for correspondence author:

Joan Webster

Centre for Clinical Nursing

Level 2, Building 34

Royal Brisbane and Women's Hospital

HERSTON, QLD 4029

Email address for corresponding author: joan_webster@health.qld.gov.au

Acknowledgement: The study was funded by a 'Queensland Nursing Council' grant and a 'Strengthening Aged Care' grant.

ABSTRACT

Objective: To compare the effectiveness of the STRATIFY falls tool with nurses clinical judgment in predicting patient falls.

Study design and setting: A prospective cohort study was conducted amongst in-patient of an acute tertiary hospital. Participants were patients over 65 years admitted to any hospital unit. Sensitivity, specificity and positive and negative predictive values of the instrument and nurses clinical judgement in predicting falls were calculated.

Results: 788 patients were screened and followed up during the study period. Falls prevalence was 9.2%. Of the 335 patients classified as being 'at risk' for falling using the STRATIFY tool, 59 (17.6%) did sustain a fall (sensitivity .82, specificity .61, positive predictive value .18, negative predictive value .97). Nurses' judged that 501 patients were at risk of falling and, of these, 60 (12.0%) fell (sensitivity .84, specificity .38, positive predictive value .12, negative predictive value .96). The STRATIFY correctly identified significantly more of the patients as either fallers or non-fallers than nurses ($p = 0.027$).

Conclusion: Considering the poor specificity and high rates of false positive results for both the STRATIFY tool and nurse's clinical judgement, we conclude that neither of these approaches are useful for screening for falls in acute hospital settings.

Key words: Accidental falls; sensitivity and specificity; positive and negative predictive value; aged; hospitals, teaching.

Running title: Screening or clinical judgement for falls risk

Word Count: Abstract 200 words; body text 2040.

What is new?

Key findings	<ul style="list-style-type: none">- Nurses clinical judgement was no better than the STRATIFY falls screening tool in predicting which patients would fall- The proportion of false positive results was unacceptably high, using either nurses' clinical judgement or the STRATIFY tool.
What this adds to what was known	<ul style="list-style-type: none">- This is the first study to test nurses clinical judgement and the STRATIFY tool in acute tertiary facility- It is the largest study to date but matches results from other centres
What is the implication, what should change now	<ul style="list-style-type: none">- Patients in acute facilities should not be routinely screened for their risk of falling using the STRATIFY falls screening tool. It remains unclear if other tools may be useful.- Future research should focus on evaluating interventions that claim to prevent falls.

INTRODUCTION

Patient falls are not uncommon in hospital settings with prevalence estimates between 0.02% and 12.2% [1], [2], and [3]. Most falls are not serious but a small proportion; 1 - 4% result in serious injury [4], [5], and [6]. Because of this, many hospitals have implemented routine screening to assess a patient's risk of falling, followed up with a more focused assessment of those deemed to be at 'high risk'. [7], [8], and [9]. One of the problems with this approach is that screening tools may have high false positive rates, over estimating the number of patients at high risk of falling. For example the false positive rates for the STRATIFY screening tool range between 37% and 85% [7], [10], [11], [12], [13], [14] and [15]. Such over estimation may lead to valuable resources being squandered on those who will not fall. Or, more seriously, screening results may be ignored completely because of the impossibility of following up on the large number of patients deemed to be at high risk.

An alternative to using a falls screening tool is to rely on clinical judgement. Three studies have assessed nurse's clinical judgement in predicting which patients may fall. The first of these compared the Morse Fall Scale and the Functional Reach test with nurse's clinical judgement [16]. Similar results were reported irrespective of the assessment method used and the investigators concluded that standardized tests were no better than nurse's clinical judgement in predicting falls. The second study was conducted in two aged care and rehabilitation wards in an acute tertiary hospital in Australia [17]. Two falls assessment tools were compared with nurse's clinical judgement. Each methods showed good sensitivity but had low positive predictive values; that is they were performed poorly in terms of their ability to accurately discriminate between those with and without a high risk of falling. In the most recent

study, the Downtown and the STRATIFY falls risk assessment tools were compared with clinical judgement (based on observation of wandering behaviour). Although clinical observation was shown to have a higher overall accuracy than either of the screening tools, none of the strategies could be recommended due to their low predictive capacity [18].

In Australia, the STRATIFY falls tool has been recommended for routine use in acute hospital settings [19] Two studies have shown that nurses' clinical judgement is as effective as using a risk assessment instrument in predicting which patients will fall. However, neither of these studies included the STRATIFY tool, nor has the STRATIFY tool been tested in an Australian setting. Consequently, there was an urgent need establish clear evidence of benefit, before widespread introduction of the screening tool. The purpose of the present study was to test the STRATIFY falls risk assessment tool in an acute hospital setting and to compare screening outcomes with nurses clinical judgment.

METHODS

Population

We included in-patients from a 982 bed general, tertiary referral teaching hospital admitted to internal medical, surgical, orthopaedic, psychiatric, oncology or geriatric rehabilitation services. We obtained Human Research Ethics approval to extract follow-up data from patient records.

Instrument

The STRATIFY tool consists of five items assessing: whether a patient was admitted with a fall or has fallen since admission, if they are agitated, have visual impairment, require frequent toileting, or have high transfer/mobility needs using the Barthel scoring system. Each item is scored with a 'yes' or 'no' and equal value is assigned to each risk. 'Yes' is scored as 1 and 'no' as 0. The authors suggest a score of equal to or greater than 2 be used as a cut off score to determine risk of falling [10]. Apart from the first item, no clear definitions are provided.

Design

We used a prospective cohort study to compare the ability of the STRATIFY tool to predict a patient's risk of falling with nurses clinical judgment.

Procedure

All patients 65 years and over were assessed for their risk of falling within 48 hours of admission by research officers trained in the use of the STRATIFY tool. Information was gathered either from patient notes or from the nurse caring for the patient. At the same time, the nurse caring for the patient was asked if they thought the patient was at risk of falling. The nurse was unaware of the STRATIFY score at this time; the response was documented as 'yes' or 'no'. To establish whether a patient fell during their hospital stay we reviewed weekly printouts from the hospital's 'Patients Incidence Reports' data base and, on hospital discharge, the medical record of each patient was examined for evidence of a fall.

Outcome

We used the World Health Organisation definition: ‘A fall is an event which results in a person coming to rest inadvertently on the ground or floor or other lower level’ for the primary outcome or ‘gold standard’ [20].

Analysis

Data were entered and analysed using SPSS version 15.0. Descriptive statistics (frequencies or means and standard deviations SDs) were used for baseline demographic and risk factor characteristics. The sensitivity (ability to correctly classify those who fall), specificity (ability to correctly classify those who do not fall), positive predictive value (ability to predict those who will fall) and negative predictive value (ability to predict those who will not fall) [21] of the STRATIFY tool and nurses’ clinical judgement were calculated using a two by two table for each test. The recommended cut-off point of ≤ 2 was used for the STRATIFY tool analysis.

RESULTS

Baseline and risk factor data are based on 801 patients who were screened between 17th March and 24th October 2007. Follow-up data was unavailable for 13 patients; so outcome data was derived from the remaining 788 (98.4%) patients. Of those screened 348 (43.5%) were medical patients, 306 (38.2%) were surgical, 55 (6.9%) were from oncology wards, 46 (5.7%) were patients in the extended stay or geriatric assessment and rehabilitation units, 37 (4.6%) were mental health patients and the remaining nine (1.1%) were from our infectious diseases unit. The mean age of the sample was 77.7 years [standard deviation (SD) 7.89] and the mean length of stay was 27.7 days (range 1 – 224 days; SD 31.74). A total of 394 (49.2%) participants were male, 260 (32.5%) had experienced a previous fall, 178 (22.2%) were classified

as 'agitated', 154 (19.2%) were visually impaired and 235 (29.3%) required frequent toileting. According to the Barthel scoring system, 312 (39.0%) patients had a transfer/mobility risk. Using the STRATIFY screening tool, 335 (42.5%) patients were classified as being at risk for falling. On ten occasions, nurses' estimate of the patients risk for falling was not obtained. Of the remaining 778 patients, nurses' judged 501 (64.4%) patients as a falls risk.

Primary outcome

Of the 335 patients classified as being 'at risk' for falling using the STRATIFY tool, 59 (17.6%) did sustain a fall, the false positive rate was 82.4% and the false negative rate was 18.1%. Adjusting the cut off point upwards, decreased sensitivity but increased specificity and the positive predictive value. Nurses' judged that 501 patients were at risk of falling and, of these, 60 (12.0%) fell. The false positive rate for nurse's clinical judgement was 88.0% but the false negative rate fell to 4.0%. Table 1 shows our data against results from other validity studies. The STRATIFY correctly identified significantly more of the patients as either fallers or non-fallers than nurses ($p = 0.027$).

DISCUSSION.

Falling is the one of the most frequently reported adverse events in hospital settings and is associated with significant morbidity [6] and health care costs [22]. Despite being of such concern, it is often difficult to identify which patients will fall. Because of this, screening for risk factors associated with falling has become commonplace. The ideal screening situation involves the use of an inexpensive, non-invasive, easily reproducible test that has a high level of sensitivity, specificity, and predictive value

and that will detect a common problem that can be treated and that if left untreated results in considerable morbidity and mortality [23]. Fall prevention is an area in which screening is clearly justified; however, the screening test of choice remains debatable.

Although the STRATIFY tool is widely used for this purpose, there is conflicting evidence about its validity [7], [10], [11], [12], [13], [14] and [15]. While the studies conducted by the author of the instrument concluded that the STRATIFY was a clinically useful predictive tool; [10] subsequent studies have suggested that original results may have been overly optimistic and have advised caution with its use [7], [11], [12], [13], [14] and [15]. Differences in the clinical characteristics of patients may explain these differences, although most studies have drawn their samples from similar settings. The results of our study reinforce the observation that the STRATIFY tool is a poor predictor of falling; because, although the sensitivity was reasonable at 82%, the majority of patients with a positive screen did not fall. Results may have been improved if we had weighted risks in terms of their predictive value [24] but we were keen to test the falls tool in the form recommended for use in our hospital by the Australian Council for Safety and Quality in Health Care [19].

Nurses predicted fewer of the patients who did fall (12%) compared with the STRATIFY tool (18%) but their clinical judgment of falls risk showed similar sensitivity to the falls screening tool. However, because nurses classified fewer patients as low risk than the STRATIFY tool, the specificity was unacceptably low at 38%. Table 1 shows that our findings are congruent with both of the previous studies of clinical judgement, where patient's risks were overestimated, leading to an inability

to effectively discriminate between those who would and those who would not fall [16] and [17]. Perhaps the clinical judgement of nurses would have been improved if the question had been “do you think this patient needs ‘one to one nursing’ to prevent a fall?”

Poor predictability of both the STRATIFY tool and nurses clinical judgement is probably best explained by the high incidence of co-morbidities amongst hospital inpatients, placing many in the high risk category; combined with a relatively low falls rate, around 9%. As positive and negative predictive values are directly related to the prevalence of the problem in the cohort [22] it is perhaps not surprising that the instrument performs better in populations where more falls occur.

This of course raises the question about whether screening should be conducted at all where falls rates are low, given the potential costs consequences of providing a false positive screen. For example, categorising the patient as ‘at risk’, either by a falls screening tool or by nurse’s clinical judgement, indicates that a full assessment by other members of the health care team should follow [19]. While screening itself is relatively cheap, five minutes or so of nursing or occupational therapist’s time, follow-up assessment and is more costly, involving, at least, a medical review and generally additional assessment by physiotherapists, dieticians and the like.

In addition, most falls are of a minor nature. Unfortunately we did not classify falls outcome as minor or major this study. However, Vassallo et al (2005) noted that, of those who fell in rehabilitation wards, 30% sustained an injury and only 5 (2.1%) of these were serious [25]. If we had included only those falls requiring treatment, the

false positive rate would have been even higher. This is not to say that we should stop trying to identify those at risk for falling in acute settings but to suggest that a different approach is needed. This may involve a second low-cost appraisal, between the initial screen and a full assessment. Perhaps a completely new instrument for acute settings based on contemporary risk factors [15] and root cause analyses is required. It is also perhaps timely to undertake a rigorous cost-effectiveness analysis of routine screening for falls risk in acute settings. This should involve measures such as the cost of screening; assessment costs; interventions implemented; additional length of hospital stay; and potential savings from avoiding injury or death. Standard instruments such as ‘disability-adjusted life years’[26] and ‘health-related quality of life’[27] should be used to measure outcomes.

LIMITATIONS

There were several limitations to this study. First, participants in the study were heterogeneous, drawn from all hospital wards. This may have affected the accuracy of the instrument but this seems unlikely. The tool asks the rater to simply identify whether certain risks are present; so answers should not be affected by the ward in which the patient is nursed. Second, the nurse caring for the patient at the time was asked, in his or her opinion, if the patient was at risk of falling. We regularly use relief staff so it is possible that the nurse was unfamiliar with the patient’s true risk level. In addition we did not conduct any formal inter-rater reliability testing among the four research nurses who collected data. However, there was extensive discussion and agreement about the meaning of items on the questionnaire before it was used, so we do not expect that this would have affected outcomes. Finally, there is also a possibility that not all falls were recorded in either the patients’ medical record or on

the hospitals adverse event reporting system, affecting the accuracy of both screening methods.

CONCLUSION

Considering the poor specificity and high rates of false positive results for both the STRATIFY tool and nurse's clinical judgement, we conclude that neither of these approaches are useful for screening for falls in acute hospital settings.

REFERENCES

1. Haflon P, Eggli Y, Van Melle G, Vagnair A. Risk of falls for hospitalized patients: A predictive model based on routinely available data. *J Clin Epidemiol* 2001;54:1258-1266.
2. Passaro A, Volpato S, Romagnoni F, Manzoli N, Zuliani G, Fellin R et al. Benzodiazepines with different half-life and falling in a hospitalized population: The GIFA study. *J Clin Epidemiol* 2000;53:1222-1229.
3. Schwendimann R, De Geest S, Milisen K. Evaluation of the Morse Fall Scale in hospitalised patients. *Age Ageing* 2006;35:311-313.
4. Schwendimann R, Buhler H, De Geest S, Milisen K. Falls and consequent injuries in hospitalized patients: effects of an interdisciplinary falls prevention program. 2006; <http://www.biomedcentral.com/1472-6963/6/69>
5. Fonda D, Cook J, Sandler V, Bailey M. Sustained reduction in serious fall-related injuries in older people in hospital. *Med J Aust* 2006;184:379-382.
6. Vassallo M, Vignaraja R, Sharma JC, Briggs R, Allen S.. The relationship of falls to injury among hospital in-patients. *Int J Clin Pract* 2005;59:17-20.
7. Coker E, Oliver D. Evaluation of the STRATIFY falls prediction tool on a geriatric unit. *Outcomes Manag* 2003;7:8-14.
8. Gowdy M, Godfrey S. Using tools to assess and prevent inpatient falls. *Jt Comm J Qual Saf* 2003;29:363-368.
9. Dempsey J. Falls prevention revisited: a call for a new approach. *J Clin Nurs* 2004;13:479-485.
10. Oliver D, Britton M, Seed P, Martin FC, Hopper AH. Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict

- which elderly inpatients will fall: case-control and cohort studies. *BMJ* 1997;315:1049-1053.
11. Chiari P, Mosci D, Fontana S. Evaluation of 2 tools for measuring the risk of falls among patients. *Assist Infirm Ric* 2002;21:117-124.
 12. Wijnia JW, Ooms ME, van Balen R. Validity of the STRATIFY risk score of falls in nursing homes. *Prev Med* 2006;42:154-157.
 13. Vassallo M, Stockdale R, Sharma JC, Briggs R, Allen S. A comparative study of the use of four fall risk assessment tools on acute medical wards. *J Am Geriatr Soc* 2005;53:1034-1038.
 14. Smith J, Forster A, Young J. Use of the 'STRATIFY' falls risk assessment in patients recovering from acute stroke. *Age Ageing* 2006;35:138-143.
 15. Webster J, Courtney M, O'Rourke P, Marsh N, Gale C, Abbott B, McRae P, Mason K. Should elderly patients be screened for their 'falls risk'? Validity of a falls screening tool and predictors of falls in a large acute hospital. *Age and Ageing*, 2008 Aug 13; doi:10.1093/ ageing/afn153
 16. Eagle DJ, Salama S, Whitman D, Evans LA., HO E, Olde J. Comparison of three instruments in predicting accidental falls in selected inpatients in a general teaching hospital. *J Gerontol Nurs* 1999;25:40-45.
 17. Myers H, Nikoletti S. Fall risk assessment: a prospective investigation of nurses' clinical judgement and risk assessment tools in predicting patient falls. *Int J Nurs Pract* 2003;9:158-165.
 18. Vassallo M, Poynter L, Sharma JC, Kwan J, Allen SC. Fall risk-assessment tools compared with clinical judgment: an evaluation in a rehabilitation ward. *Age Ageing* 2008;37:277-81

19. Australian Council for Safety and Quality in Health Care. Preventing falls and harm from falls. Best practice guidelines for Australian hospitals and residential aged care facilities. Canberra: Safety & Quality Council; 2005.
20. World Health Organization. Injuries and violence prevention. Available at www.who.int/violence_injury_prevention/unintentional_injuries/falls/falls1/en / (accessed 1 December 2007).
21. Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R. Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol 2008;56:45-50.
22. Hendrie D, Hall SE, Arena G, Legge M. Health system costs of falls of older adults in Western Australia. Aust Health Rev 2004;28:363-373.
23. Nielsen C, Lang RS. Principles of screening. Med Clin North Am 1999; 83:1323-1237.
24. 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 Jan 21;2:1.
25. Vassallo M, Vignaraja R, Sharma JC, Briggs R, Allen S. The relationship of falls to injury among hospital in-patients. Int J Clin Pract 2005;59:17- 20.
26. Murray CJ, Acharya AK. Understanding DALYs (disability-adjusted life years). J Health Econ 1997;16:703-730.
27. Khanna D, Tsevat J. Health-related quality of life--an introduction. Am J Manag Care 2007;9:S218-23.

Table 1. Validity of falls assessment methods across studies

Study	Year	Falls rate	Setting	Mean age	Sensitivity	Specificity	PPV*	NPV†
Eagle et al ¹⁸ (n = 98)	1999	30.6	Aged care & rehabilitation wards	69.0				
Functional reach test					.76	.34	.33	.77
Morse Falls Scale					.72	.51	.38	.81
Nurses clinical judgement					.76	.49	.39	.83
Myers & Nikolett ¹⁹ (n = 226)	2003	15.0	Aged care & rehabilitation wards	84.9				
Falls assessment tool 1					.91	.25	.18	.94
Falls assessment tool 2					.91	.27	.18	.94
Nurses clinical judgement					.88	.26	.18	.92
Vassallo et al (n = 200)	2008	25.5	Geriatric rehabilitation hospital	80.9				
Downtown					.92	.36	.33	.93
STRATIFY					.82	.34	.30	.85
Clinical judgement (wandering)					.43	.91	.61	82.3

Webster (n = 783)	2008	9.2	Acute tertiary hospital	77.7				
STRATIFY								
Using 1 as the cut point					.93	.29	.12	.98
Using 2 as the cut point					.82	.61	.18	.97
Using 3 as the cut point					.45	.83	.22	.94
Using 4 as the cut point					.25	.95	.35	.93
Nurses clinical judgement					.85	.38	.12	.96

* Positive predictive value

† Negative predictive value