MonashHealth

Clinicians, Consumers, Evidence
Centre for Clinical Effectiveness

Rapid Review

Accuracy of validated falls risk assessment tools and clinical judgement

Citation Yap G. and Melder A. Accuracy of validated falls risk assessment tools and clinical judgement: Rapid Review. Centre for Clinical Effectiveness, Monash Innovation and Quality, Monash Health, Melbourne, Australia.

Abstract

Background

The Executive Director Residential and Support Services/Chief Nursing and Midwifery Officer has requested that the Centre for Clinical Effectiveness undertake a review of literature for evidence that suggests that the use of validated risk scoring tools may reduce the use of clinical judgement, and if there more effective alternatives. Monash Health currently uses a modified stratified Ontario risk assessment tool.

Objective

1) To find out if there is evidence to show that risk assessment tools decrease the use of clinical judgement.

2) To identify the best validated risk assessment tools (or alternative) that accurately assesses and predict falls in hospital inpatients.

Methodology

A search for synthesised evidence published in English from 2012 to present was conducted in Ovid MEDLINE® and CINAHL Plus (Table 5 and Table 6) and documents were screened according to inclusion criteria and exclusion criteria listed in Table 2. A brief assessment was made on the quality of the evidence included in the report based on selected criteria (Page 4). Risk assessment tools, screening tools or prediction tools are included in the review, and accuracy of a risk assessment tool is represented by its sensitivity and specificity values.

Results

A total of eight documents were selected for inclusion: four high quality reviews/meta-analyses evaluating validated fall risk assessment tools that were published in the acute aged [1], paediatric [2], inpatient rehabilitation [3], and adult inpatient settings [4] (Table 4); three prospective studies published in the aforementioned settings [5-7] (Table 4); and one evidence-based guideline published by the National Institute For Health and Care Excellence (NICE) [8].

1) From the evidence identified in this review, there was none to suggest that the use of fall risk assessment tools reduced the use of clinical judgement.

2) No risk assessment tools in predicting falls in inpatients have been found to be reliable and valid across institutions and populations in the paediatric setting [2]. Although there is evidence of STRATIFY (a validated tool for assessing risk of falls) demonstrating poor sensitivity [5], specificity [7], and predictive accuracy [1], synthesised results from meta-analyses revealed that STRATIFY was the best tool in assessing risk of falls among acute adult inpatients (>16 years) [4]; as well as in older inpatients (≥ 65 years) across hospital settings [1]. Neither self-reporting nor clinical judgement were recommended in predicting falls in a geriatric rehabilitation hospital [6] or across specialty wards in hospitals [3].

Table1. Summary of the prediction values (sensitivity, specificity, PPV, NPV) of validated risk assessment tools reported in the meta-analyses included in the report.

Tool	Setting	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Reference
STRATIFY	Acute; Adults > 65 y	63% (54–69%)	71% (67–73%)	NR	NR	[1]
	Acute; Adults > 16 y	80% (72–86%)	68% (66-69%)	NR	NR	[4]
	Rehabilitation; Adults ≥ 65 y	80% (72–86%)	68% (66-69%)	NR	NR	[3]
HFRM-II	Acute; Adults > 65 y	92% (84–97%)	37% (33–41%)	NR	NR	[1]
	Acute; Adults > 16 y	63% (55–70%)	64% (63–65%)	NR	NR	[4]
MFS	Acute; Adults > 16 y	76% (70–80%)	68% (66–70%)	NR	NR	[4]
Conley	Acute; Adults > 65 y	69%	41%	NR	NR	[1]
Downton	Rehabilitation; Adults ≥ 65 y	92% (82–97%)	36% (28–43%)	33% (25–41%)	93% (83–97%)	[3]
Clinical Judgement	Rehabilitation; Adults ≥ 65 y	43% (30–56%)	91% (84–94%)	61% (44–75%)	82% (75–87%)	[3]

PPV – positive prediction value; NPV – negative prediction value; CI – confidence interval; y – years; NR – not reported;

The accuracy of a risk assessment tool is represented by its sensitivity and specificity (see Page 3 – Definitions). A fall risk prediction tool may be considered to be of high predictive value when it demonstrates a sensitivity and specificity above 70% [8]. Table 3 (Page 5) lists the full names of the risk assessment tools mentioned above.

Conclusion

In conclusion, the body of evidence indicates that tools with a numerical predictor of risk should no longer be used; however a multifactorial risk assessment followed by multifactorial interventions tailored to the patient's needs is recommended. [8] Evidence confirms a general consensus that fall risk assessment tools should never replace clinical judgement [1, 2, 7], but both to be integrated into a multifactorial intervention strategy. [1, 7] Although it may be a preferred option to adopt the same fall risk screening tool across hospital units, selecting an appropriate risk assessment tool that takes into account individual risk factors is recommended by several researchers [1, 2, 5].

Rapid Review – full report

Accuracy of validated falls risk assessment tools and clinical judgement

Citation Yap G. and Melder A. 2017. Accuracy of validated falls risk assessment tools and clinical judgement: Rapid Review. Centre for Clinical Effectiveness, Monash Innovation and Quality, Monash Health, Melbourne, Australia.

Background

The Executive Director Residential and Support Services/Chief Nursing and Midwifery Officer has requested that the Centre for Clinical Effectiveness undertake a review of literature for evidence that suggests that the use of validated risk scoring tools may reduce the use of clinical judgement, and if there more effective alternatives. Monash Health currently uses a modified stratified Ontario risk assessment tool.

Question(s)

1) Does the use of risk assessment tools decrease the use of clinical judgement?

2) What are the validated risk assessment tools that accurately predict falls in hospital inpatients?

Definitions

For the purpose of this review, risk assessment tools will refer to tools used to accurately predict the risk or likelihood of falls in hospitalised patients. These include risk assessment tools, screening tools or prediction tools as these terms are used interchangeably in literature.

The accuracy of a risk assessment tool is represented by its sensitivity and specificity. From published literature, sensitivity and specificity are the most appropriate and common measures for prediction [8] and fall risk prediction tools may be considered to be of high predictive value when they demonstrate sensitivities and specificities above 70% [8].

In this report, sensitivity is a measure of how well a tool correctly classifies or predicts a patient at high risk to fall and the patient fell; specificity is a measure of how well a tool correctly classifies the patient as low risk to fall and the patient did not fall [2]. The positive predictive value (PPV) is the probability the patient who scores as high risk will actually fall; while the negative predictive value (NPV) is the probability the patient who scores as low risk will not fall. Therefore a high PPV indicates likelihood a person with a high risk score will fall [2].

Search Methods

Inclusion and exclusion criteria

Table 2. Inclusion/exclusion criteria

Setting	*Include:* Hospitals (public and private), international; inpatients (paediatric and adult) across specialties *Exclude:* Home care; rehabilitation; residential aged care; community dwelling; nursing homes
Intervention	*Include*: Evaluation of validated fall risk assessment tools, other strategies of assessing fall risk (including clinical judgement) *Exclude*: Evaluation of any other intervention or combination of strategies; non-validated fall risk assessment tools
Outcome	Accuracy of prediction of falls • Sensitivity • Specificity • Positive and negative predictive values (PPV and NPV), where reported
Publications details	*Inclusion*: Systematic reviews, meta-analysis; prospective studies evaluating two or more fall risk assessment tools *Exclusion*: Reviews or comparative studies evaluating a single tool
Limitation	2012-2017; Published in English; Humans
Databases	Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE and Versions(R); CINAHL Plus

Document Selection

The search methodology (Table 2) was decided *apriori* and searches were performed in two medical databases according to the inclusion and exclusion criteria. Only the most updated synthesised data (i.e., reviews or meta-analysis) were selected; and comparative studies published only after the date (i.e., year) of the most recent systematic review were selected. References were exported and screened in Endnote X7 (Thompson, Reuters, Carlsbad, California, USA) by one reviewer (GY).

Assessment of Quality

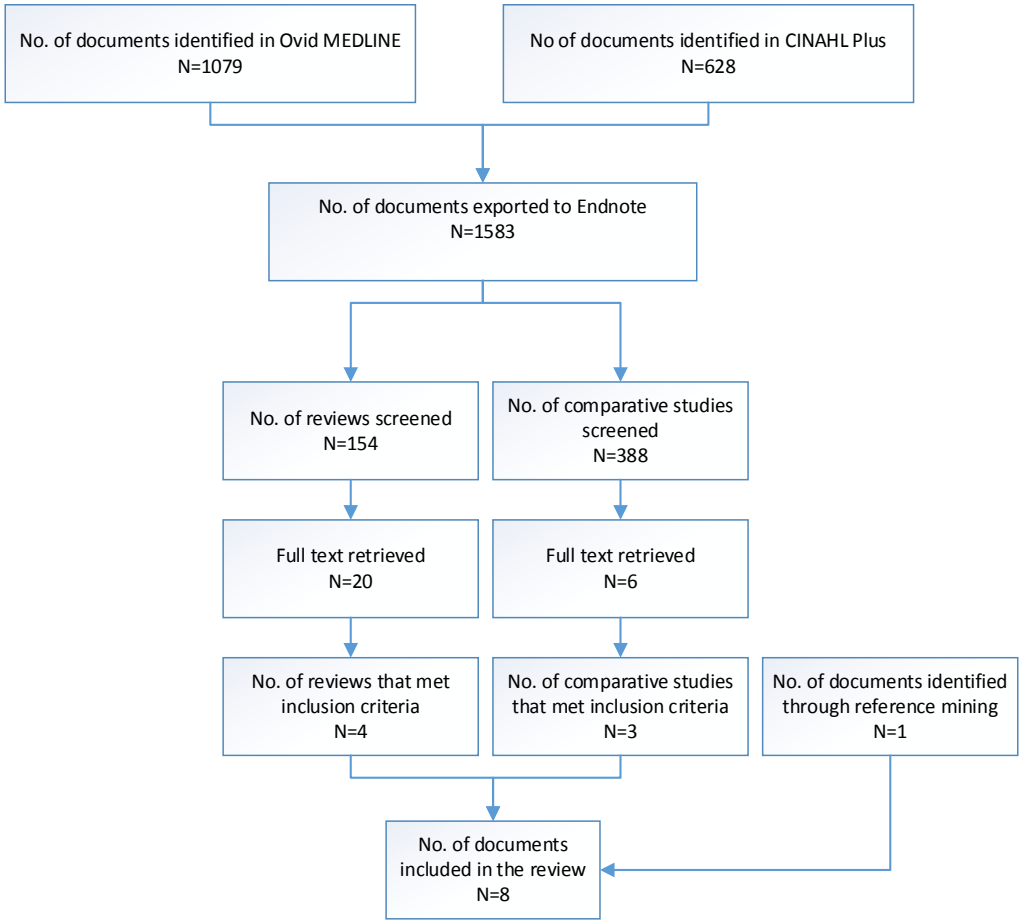
A brief assessment was made on the quality of the included reviews in the report based on three criteria: 1) A clear search strategy detailing the search terms, and date limitations was included; 2) Inclusion and exclusion criteria was detailed; 3) An appraisal of the quality of included studies was conducted. Reviews were ranked as high quality if they met all three criteria, moderate quality if they met only two criteria, and low quality if they met less than two of the criteria.

An assessment was made on the quality of the included comparative studies based on the ten criteria listed in a methodological evaluation used in an included systematic review [4]. Studies were considered to be of high quality if they achieved seven or more criteria; moderate quality if only five to six criteria were met; and low quality if four or less criteria were met.

The quality of included reviews and studies are presented in Table 4.

Search Results

Figure 1. Flow diagram of the search results.



Database search

The databases search resulted in 1583 documents which were exported to Endnote.

Screening

Using the search function within Endnote, 154 reviews were identified for screening of title and abstract. Of the twenty full texts retrieved, four reviews met the inclusion criteria for selection.

Document selection

Of the four articles selected, the most recent review (in adult inpatient setting) was published in 2015; hence comparative studies published in 2016 and 2017 were selected in accordance to the inclusion criteria listed in Table 2.

Three comparative studies met the inclusion criteria for selection.

Reference mining of full-text reviews identified one further document that met the inclusion criteria for selection.

Description of included studies

A total of eight documents were selected for inclusion: four high quality reviews/meta-analysis evaluating validated fall risk assessment tools that were published in the acute aged [1], paediatric [2], inpatient rehabilitation [3], and adult inpatient settings [4] (Table 4); three recent prospective studies published in the aforementioned settings [5-7] (Table 4); and one evidence-based guideline published by the National Institute For Health and Care Excellence (NICE) [8].

There was no synthesised evidence (i.e., from systematic reviews or meta-analysis) that specifically reported the impact of fall risk assessment tools on clinical judgement. Evidence that compared the use of clinical judgement to the use of validated risk assessment tools were reported in the included studies of one review [3]. Only one recent prospective study was identified that compared the accuracy of STRATIFY, self-reporting and clinical judgement [6].

A summary of the predictive accuracy of validated risk assessment tools (sensitivity and specificity; positive and negative predictive values), and the use of clinical judgement in different inpatient settings are presented in Table 4.

Table 3. List of risk assessment tools presented in the report and their commonly-used abbreviations.

STRATIFY	St. Thomas Risk Assessment Tool in Falling elderly inpatients	***CHAMPS***	Change in mental status, History of falls, Age <36 months, Mobility impairment, Parental involvement and Safety
DOWNTON	Downton Fall Risk Index	***GRAF-PIF***	General Risk Assessment For Paediatric In-patient Falls
HRFM-II	Hendrich II Fall Risk Model	***HDFS***	Humpty Dumpty Fall Scale
OM	Ontario Modified STRATIFY	***I'm SAFE***	Impairment, Medications, Sedation/anesthesia, Admitting diagnosis, Fall history, and Environment of care
TNF	The Northern Hospital Modified STRATIFY	***PJC-FRAT***	Peter James Centre Falls Risk Assessment Tool
MFS	Morse Falls Scale		

Synthesis of Results

The use of clinical judgement in risk assessment for falls

Inpatient rehabilitation

The authors of a review of risk assessment tools in an inpatient rehabilitation setting reported findings from an included study [3]. Evidence from the study by Vassallo and colleagues (2008), showed that the validated falls risk assessment tools, STRATIFY and DOWNTON, had better sensitivity (82% and 92%, respectively) than clinical judgment (43%), but both had worse specificity (34% and 36%, respectively) than clinical judgment (91%); hence the authors questioned the usefulness of the two falls prediction tools [3]. In a recent prospective study that compared the accuracy of STRATIFY, self-reporting and clinical judgement in predicting falls in a geriatric rehabilitation hospital, overall, clinical judgement demonstrated the second highest sensitivity at admission (67%) after self-reporting assessment (71%), but the lowest specificity (46%) as compared to STRATIFY (68%) and self-reporting (65%) [6]. Nevertheless, none of the three techniques can be recommended due to low predictive value <70% [6].

Inpatients across specialties

Strupeit *et al.* (2016) mentioned that in a study of STRATIFY and nurses' clinical judgment in multiple hospital wards in inpatient specialties, Webster and colleagues (2010) recommended against using either STRATIFY or nurses' clinical judgment in screening for falls in acute inpatient settings due to a lack of demonstrable efficacy [6].

We were unable to find evidence suggesting that the use of fall risk assessment tools reduced the use of clinical judgement.

Accuracy of prediction of falls using validated risk assessment tools

Paediatric setting

Evidence from one high quality literature review published in 2017 suggests that among the several paediatric fall risk assessment tools published in the literature (i.e., CHAMPS, Cummings Scale, GRAF-PIF, HDFS, I'm SAFE), none have been found to be reliable and valid across institutions and diverse populations [2].

Acute-care inpatient populations

In adults > 16 years: Another 2013 systematic review and meta-analysis that included evidence from prospective validation studies, showed the STRATIFY tool to be the best tool for assessing the risk of falls among hospitalised acutely ill adult patients, followed by MFS and finally HFRM-II [4]. Results from a recent prospective multicentre study that compared DOWNTON and STRATIFY revealed that that the two instruments had poor sensitivity values, according to the optimal cut-off value that was identified (STRATIFY: 47.6%; DOWNTON: 66.7%). As such, the two falls risk assessment instruments have proved to be of limited utility as tools for detecting the risk of falling among acute care inpatients [5].

In aged population ≥ 65 years: Based on evidence from prospective validation studies on inpatients > 65 years across hospital settings, a 2015 systematic review and meta-analysis revealed that the HFRM-II obtained higher sensitivity value than STRATIFY (92% vs. 63%), while the STRATIFY showed higher specificity than HFRM-II (71% vs. 37%). However, the authors conclude that there was a low predictive accuracy for both tools [1]. A prospective study comparing the OM, TNH and STRATIFY showed that the OM had significantly worse accuracy compared to TNH and STRATIFY, with no significant differences in accuracy between TNF and STRATIFY [7].

Inpatient rehabilitation

A 2012 systematic review and meta-analysis based on evidence from prospective investigative studies (with major methodological limitations) indicated that the DOWNTON tool had the highest sensitivity (92%), while the PJC-FRAT offered the best balance between sensitivity and specificity (73% and 75%, respectively) [3]. Neither STRATIFY, self-reporting or clinical judgement were recommended in predicting falls in a geriatric rehabilitation hospital due to low predictive value of <70% [6].

NICE guidance

NICE reported that although the falls risk prediction tools reviewed demonstrated sensitivities and specificities above the desired threshold (>70%), none of the tools were adequately replicated in the relevant setting, and their validity and reliability are uncertain. The quality of the evidence presented was low or very low. [8] The most recent NICE guidance (2013) states that tools with a numerical predictor of risk should no longer be used; however, multifactorial risk assessment followed by multifactorial interventions tailored to the patient's needs is recommended. [8].

Conclusions

There was no evidence to suggest that the use of fall risk assessment tools reduced the use of clinical judgement.

No risk assessment tools have been found to be reliable and valid across institutions and diverse populations in the paediatric setting [2]. Although there is evidence of STRATIFY demonstrating poor sensitivity [5], specificity [7], and predictive accuracy [1], synthesised results from meta-analyses revealed that STRATIFY was the best validated tool in assessing risk of falls among acute adult inpatients (>16 years) [4]; as well as in older inpatients (≥ 65 years) across hospital settings [1]. The PJC-FRAT offers the best balance of sensitivity and specificity (>70%) in an inpatient rehabilitation population [3].

Neither self-reporting nor clinical judgement were recommended in predicting falls in a geriatric rehabilitation hospital [6] or across specialty wards in hospitals [3], due to their low predictive value <70% [6]

In conclusion, the body of evidence indicates that tools with a numerical predictor of risk should no longer be used; however a multifactorial risk assessment followed by multifactorial interventions tailored to the patient's needs is recommended. [8] Evidence confirms a general consensus that fall risk assessment tools should never replace clinical judgement [1, 2, 7], but both to be integrated into a multifactorial intervention strategy. [1, 7] Although it may be a preferred option to adopt the same fall risk screening tool across hospital units, selecting an appropriate risk assessment tool that takes into account individual risk factors is recommended by several researchers [1, 2, 5].

Limitations

It is worth noting that the reviews and meta-analyses included in this report used different methods and scales for rating the quality of their included studies [1-4].

Although the documents selected in the report only included data from prospective validation studies, the observational design for some of the included studies may lead authors to infer conclusions that do not have the clear cause-effect relation that is characteristic of experimental study designs, and residual confounding variables may be associated with the results [4].

Synthesis of Results

Table 4. Summary of the accuracy (sensitivity and specificity; positive and negative predictive values) of validated risk assessment tools in different inpatient settings.

Reference	Document Type	Quality	Inpatient Setting	Tools Evaluated	Accuracy (95% CI)	Results
DiGerolamo *et. al.* (2017) [2]	Review	High; Included good quality studies	Paediatric	CHAMPS paediatric FRAT	Individual data not reported	Though there are several paediatric FRATs published in the literature, none have been found to be reliable and valid across institutions and diverse populations. Based on an included 2010 retrospective study comparing the five published tools, all but CHAMPS had reasonably acceptable values for reliability at >60%; the GRAF-PIF and Cummings were most accurate in correctly identifying patients who fell. Though errors were highest with HDFS, this was the only one of the five to correctly identify the two children who fell in the prospective sample as high risk. Another retrospective study on CHAMPS, GRAF-PIF and HDFS reported that sensitivity, specificity, and error were below standard for the HDFS and GRAF PIF; and precision and accuracy in predicting paediatric fall risk was lacking in the three tools.
				Cummings scale		
				GRAF-PIF		
				HDFS		
				I'M SAFE		
Matarese *et.al.* (2015) [1]	Systematic review Meta-analysis	High; Included good quality, homogeneous studies	Acute care (across hospital setting) Aged > 65 Years;	STRATIFY	Sensitivity: 63% (54–69%) Specificity: 71% (67–73%)	Based on evidence from prospective validation studies, the HFRM-II obtained higher sensitivity value than STRATIFY (0.92 vs. 0.61), while the STRATIFY showed higher specificity than HII-FRM (0.71 vs. 0.37). The Youden Index* was quite similar between both tools (STRATIFY = 0.29 vs. HII-FRM = 0.34) showing low predictive accuracy for both tools. Conley scale used in one prospective study (n=1620) showed a sensitivity of 69% and a specificity of 41% and poor accuracy value (Youden index = 0.10). *The Youden Index summarises the predictive accuracy of a tool into a single numeric value and gives equal weight to test errors (false positive and negative). It is calculated adding sensitivity to specificity and subtracting one. The closer the index is to one, the higher is the predictive accuracy of the tool*.
				HFRM-II	Sensitivity: 92% (84–97%) Specificity: 37% (33–41%)	
				Conley Scale	Sensitivity: 69% Specificity: 41%	

Aranda-Gallardo *et.al.* (2013) [4]	Systematic review Meta-analysis	High; Quality of included studies varied	Acute-care; Adult >16 years; Mean age 70 ± 10 years Excludes rehabilitation	STRATIFY	Sensitivity: 80% (72–86%) Specificity: 68% (66-69%)	All included studies conducted a prospective validation of the various instruments examined. The results obtained showed the STRATIFY tool to be the best tool for assessing the risk of falls among hospitalised acutely ill adult patients, followed by MFS and finally HFRM-II. The STRATIFY tool provided greater diagnostic validity, with a diagnostic odds ratio value of 7.64 (4.86 – 12.00).
				MFS	Sensitivity: 76% (70–80%) Specificity: 68% (66–70%)	
				HFRM-II	Sensitivity: 63% (55–70%) Specificity: 64% (63–65%)	
Roza da Costa et.al. (2012) [3]	Systematic review Meta-analysis	High; Quality of included studies varied	Rehabilitation; Aged ≥65 years	STRATIFY	Sensitivity: 73% (63-81%) Specificity: 42% (34-51%)	Only one tool (PJC-FRAT) that was developed and tested in an elderly population of a rehabilitation hospital. Only the STRATIFY tool was assessed in all three studies; the other identified tools (PJC-FRAT and DOWNTON) were assessed by a single study. For a STRATIFY cut-score of two, pooled sensitivity was 73% (95%CI 63 to 81%) and pooled specificity was 42% (95%CI 34 to 51%). An indirect comparison of the tools across studies indicated that the DOWNTON tool has the highest sensitivity (92%), while the PJC-FRAT offers the best balance between sensitivity and specificity (73% and 75%, respectively). All studies presented major methodological limitations.
				PJC-FRAT	Sensitivity: 73% (55–90%) Specificity: 75% (66–83%)	
				DOWNTON	Sensitivity: 92% (82–97%) Specificity: 36% (28–43%) PPV: 33% (25–41%) NPV: 93% (83–97%)	
				Clinical judgement	Sensitivity: 43% (30–56%) Specificity: 91% (84–94%) PPV: 61% (44–75%) NPV: 82% (75–87%)	
Aranda-Gallardo *et. al.* (2017) [5]	Prospective multicentre (n=1247)	Moderate	Acute-care; Adult > 16 years; Mean age 66 ± 18 years	DOWNTON	Sensitivity: 66.7% Specificity: 55.3% PPV: 6% NPV: 98%; (cut point 2)	The cut-off points of both scales are defined: STRATIFY – a score ≥ 2 indicates a “high risk of falls”; Downton – scores ≥ 3 indicate a “high risk of falls”. Mean STRATIFY score was higher among surgical patients (mean score 0.82; 95% CI: 0.77– 0.87) than in those treated in the ICU (mean score 0.23; 95% CI: 0.17–0.29; p < 0.001) and also higher than among the medical patients (mean score 0.77; 95% CI 0.74–0.81; p < 0.001). In contrast, with the Downton index, the medical patients obtained a mean score (2.72; 95% CI: 2.64–2.81) that was, significantly, 0.29 points higher than that of the surgical patients (2.44; 95% CI 2.32–2.55; p < 0.001) and 0.55 points higher than that of the ICU patients (1.89; 95% CI; 1.74–2.04; p < 0.001).
				STRATIFY	Sensitivity: 47.6% Specificity: 85% PPV: 11% NPV: 98%; (cut point 1)	

Latt *et.al.* (2016) [7]	Prospective cohort (n=217)	High	Acute care; Aged ≥ 65 years	STRATIFY (OM)	Sensitivity: 80% (59–92%) Specificity: 37% (31–44%) PPV: 11% (7–18%) NPV: 95% (88–98%)	The screening tools did not differ significantly in predictive values, although PPVs were low and NPVs high. The OM had significantly worse accuracy (41.0, CI 34.7 to 47.7%, P < 0.0001) compared to TNH (54.4, CI 47.8 to 61.0%) and STRATIFY (63.1, CI 56.5 to 69.3%). Based on calculated area under the curve (AUC) of sensitivity vs. specificity, there were no significant differences in accuracy between TNH and the STRATIFY between any of the screening tools or in AUC between any of the screening tools.
				STRATIFY (TNH)	Sensitivity: 85% (64-95%) Specificity: 51% (44–58%) PPV: 15% (10–23%) NPV: 98% (92–99%)	
				STRATIFY	Sensitivity: 80% (58–91%) Specificity: 61% (55–68%) PPV: 17% (11–26%) NPV: 97% (92–99%)	
Strupeit *et. al.* (2016) [6]	Prospective cohort (n=124)	Moderate	Rehabilitation; Aged > 60 years Mean age 83 ± 8 years	STRATIFY	Sensitivity (T1; T2): 28%; 38% Specificity (T1; T2): 68%; 69% PPV (T1; T2): 60%; 11% NPV (T1; T2): 36%; 92%	T1: at admission; T2: at 3 weeks follow-up. Overall, the highest sensitivity was demonstrated for the self-report assessment. The clinical assessment exhibited a similar sensitivity, whereas STRATIFY exhibited a relatively low sensitivity. In contrast, STRATIFY displayed the highest specificity, followed by the self-report assessment and the clinical assessment. The highest PPV and NPV were obtained with the self-report assessment. The clinical assessment displayed the next-highest PPV and NPV. STRATIFY exhibited the lowest PPV and NPV. The PPVs for all of the assessment techniques were lowest at Time 2, whereas the NPVs were above 90%. The assessment through self-report (the fear of falling) exhibited the highest sensitivity, PPVs, and NPVs. STRATIFY displayed the highest specificity but the lowest sensitivity, PPVs, and NPVs. The clinical assessment exhibited moderate sensitivity, PPVs, and NPVs but had lower specificity than the other two assessment techniques.
				Clinical Judgement	Sensitivity (T1; T2): 67%; 67% Specificity (T1; T2): 46%; 41% PPV (T1; T2): 65%; 10% NPV (T1; T2): 48%; 93%	
				Self–Report (fear of falling)	Sensitivity (T1; T2): 71%; 56% Specificity (T1; T2): 65%; 58% PPV ((T1; T2): 71%; 11% NPV (T1; T2): 51%; 93%	

NICE CG 161 (2013) [8]	Evidence-based guideline	High; Based on low quality studies	Inpatients; Aged 50 – 64 years, > 65 years.	Risk prediction tools	- Recommendation: Do not use fall risk prediction tools to predict inpatients' risk of falling in hospital. - Recommendation: Regard the following groups of inpatients as being at risk of falling in hospital and manage their care according to recommendations: • All patients aged ≥ 65 years • Patients aged 50 – 64 years who are judged by a clinician to be at higher risk of falling because of an underlying condition.
					Multifactorial falls risk assessment - Older people who present for medical attention because of a fall, or report recurrent falls in the past year, or demonstrate abnormalities of gait and/or balance should be offered a multifactorial falls risk assessment. - This assessment should be performed by a healthcare professional with appropriate skills and experience, normally in the setting of a specialist falls service. - This assessment should be part of an individualised, multifactorial intervention. Multifactorial assessment may include the following: • Identification of falls history • Assessment of gait, balance and mobility, and muscle weakness • Assessment of osteoporosis risk • Assessment of the older person's perceived functional ability and fear relating to falling • Assessment of visual impairment • Assessment of cognitive impairment and neurological examination • Assessment of urinary incontinence • Assessment of home hazards • Cardiovascular examination and medication review.

KEY: CHAMPS – Change in mental status, History of falls, Age <36 months, Mobility impairment, Parental involvement and Safety; PJC-FRAT – Peter James Centre Falls Risk Assessment Tool; GRAF-PIF – General Risk Assessment For Paediatric In-patient Falls; HDFS – Humpty Dumpty Fall Scale; I'M SAFE – Impairment, Medications, Sedation/anesthesia, Admitting diagnosis, Fall history, and Environment of care; STRATIFY – St. Thomas Risk Assessment Tool in Falling elderly inpatients; MFS – Morse Falls Scale; HFRM- II –Hendrich II Fall Risk Model; STRATIFY (OM) – Ontario Modified STRATIFY; STRATIFY (TNF) – The Northern Hospital Modified STRATIFY

Sensitivity – how well a tool correctly classifies or predicts a patient at high risk to fall and the patient fell; specificity – how well a tool correctly classifies the patient as low risk to fall and the patient did not fall. PPV – positive predictive value is the probability the patient who scores as high risk will actually fall; NPV – negative predictive value is the probability the patient who scores as low risk will not fall. A high PPV indicates likelihood a person with a high risk score will fall [2].

References

1. Matarese, M., et al., *Systematic review of fall risk screening tools for older patients in acute hospitals.* Journal of Advanced Nursing, 2015. **71**(6): p. 1198-209.
2. DiGerolamo, K. and K.F. Davis, *An Integrative Review of Paediatric Fall Risk Assessment Tools.* Journal of Paediatric Nursing, 2017. **34**: p. 23-28.
3. Roza da Costa, B., et al., *Can Falls Risk Prediction Tools Correctly Identify Fall- Prone Elderly Rehabilitation Inpatients? A Systematic Review and Meta-Analysis.* PLoS ONE [Electronic Resource], 2012. **7**(7): p. e41061.
4. Aranda-Gallardo, M., et al., *Instruments for assessing the risk of falls in acute hospitalized patients: a systematic review and meta-analysis.* BMC Health Services Research, 2013. **13**: p. 122.
5. Aranda-Gallardo, M., et al., *Diagnostic validity of the STRATIFY and Downton instruments for evaluating the risk of falls by hospitalised acute-care patients: a multicentre longitudinal study.* BMC Health Services Research, 2017. **17**: p. 1-9.
6. Strupeit, S., A. Buss, and K. Wolf-Ostermann, *Assessing Risk of Falling in Older Adults-A Comparison of Three Methods.* Worldviews on Evidence-Based Nursing, 2016. **13**(5): p. 349-355.
7. Latt, M.D., et al., *The validity of three fall risk screening tools in an acute geriatric inpatient population.* Australasian Journal on Ageing, 2016. **35**(3): p. 167-173.
8. National Institute For Health and Care Excellence (NICE) *Falls: assessment and prevention of falls in older people. 2013.* . 2013 Accessed August 2016].

Appendix 1

Table 5. Ovid MEDLINE Database search

Search terms in Ovid Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE and Versions(R)	
1	*Accidental Falls/pc [Prevention & Control]
2	Falls.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
3	*Accidental Falls/ or *Safety Management/
4	*Inpatients/
5	hospital.mp. or *Hospitals/
6	*Risk Assessment/mt, og, st, ut [Methods, Organization & Administration, Standards, Utilization]
7	*Risk Assessment/ or *Judgment/ or *Clinical Competence/
8	(predict* or strateg* or intervention* or tool*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
9	(sensitivity and specificity).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
10	1 or 2 or 3
11	4 or 5
12	6 or 7 or 8 or 9
13	10 and 11 and 12
14	limit 13 to (english language and humans and yr="2012 -Current")
	Results: n=1029

Table 6. CINAHL Plus Database search terms

Search terms in CINAHL Plus	
S1	(MH "Accidental Falls/PC")
S2	(MM "Fall Risk Assessment Tool") OR "falls" OR (MM "Fall Prevention (Iowa NIC)") OR (MM "Safety Behavior: Fall Prevention (Iowa NOC)")
S3	(MM "Inpatients") OR "inpatients"
S4	"hospital"
S5	(MM "Risk Assessment") OR "risk assessment" OR (MM "Fall Risk Assessment Tool")
S6	(MH "Decision Making, Clinical") OR (MH "Clinical Competence") OR "clinical judgement" OR (MH "Clinical Assessment Tools")
S7	"predict"
S8	"strategy"

S9	"intervention"
S10	"tool"
S11	(MM "Sensitivity and Specificity") OR "sensitivity and specificity"
S12	S1 OR S2
S13	S3 OR S4
S14	S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11
S15	S12 AND S13 AND S14
S16	*Limiters* - Publication Year: 2012-2017; Published Date: 20120101-20171231; English Language; Language: English
S17	S15 AND S16
	Results: n = 628