



## EXTERNAL VALIDATION OF APPROACHES TO PREDICTION OF FALLS DURING HOSPITAL REHABILITATION STAYS AND DEVELOPMENT OF A NEW SIMPLER TOOL

Angela VRATSISTAS-CURTO, BappSc(OT)<sup>1</sup>, Anne TIEDEMANN, PhD<sup>1</sup>, Daniel TREACY, DHM<sup>2</sup>, Stephen R. LORD, DSc<sup>3</sup> and Catherine SHERRINGTON, PhD<sup>1</sup>

From the <sup>1</sup>Musculoskeletal Health Sydney, School of Public Health, The University of Sydney, <sup>2</sup>Physiotherapy Department, Prince of Wales Hospital, and <sup>3</sup>Neuroscience Research Australia, University of New South Wales, Sydney, Australia

**Objectives:** To test the external validity of 4 approaches to fall prediction in a rehabilitation setting (Predict\_FIRST, Ontario Modified STRATIFY (OMS), physiotherapists' judgement of fall risk (PT\_Risk), and falls in the past year (Past\_Falls)), and to develop and test the validity of a simpler tool for fall prediction in rehabilitation (Predict\_CM2).

**Participants:** A total of 300 consecutively-admitted rehabilitation inpatients.

**Methods:** Prospective inception cohort study. Falls during the rehabilitation stay were monitored. Potential predictors were extracted from medical records.

**Results:** Forty-one patients (14%) fell during their rehabilitation stay. The external validity, area under the receiver operating characteristic curve (AUC), for predicting future fallers was: 0.71 (95% confidence interval (95% CI): 0.61–0.81) for OMS (Total\_Score); 0.66 (95% CI: 0.57–0.74) for Predict\_FIRST; 0.65 (95% CI 0.57–0.73) for PT\_Risk; and 0.52 for Past\_Falls (95% CI: 0.46–0.60). A simple 3-item tool (Predict\_CM2) was developed from the most predictive individual items (impaired mobility/transfer ability, impaired cognition, and male sex). The accuracy of Predict\_CM2 was 0.73 (95% CI: 0.66–0.81), comparable to OMS (Total\_Score) ( $p=0.52$ ), significantly better than Predict\_FIRST ( $p=0.04$ ), and Past\_Falls ( $p<0.001$ ), and approaching significantly better than PT\_Risk ( $p=0.09$ ).

**Conclusion:** Predict\_CM2 is a simpler screening tool with similar accuracy for predicting fallers in rehabilitation to OMS (Total\_Score) and better accuracy than Predict\_FIRST or Past\_Falls. External validation of Predict\_CM2 is required.

**Key words:** aged; accidental falls; rehabilitation clinical prediction.

Accepted Sep 28, 2017; Epub ahead of print xx, 2017

J Rehabil Med 2018; 50: 00–00

Correspondence address; Angela Vratisstas-Curto, Musculoskeletal Health Sydney, School of Public Health, The University of Sydney, Sydney, Australia. E-mail: angela.vratisstas@sydney.edu.au

Falls and fall-related injuries pose a significant burden to older people, their families and healthcare

systems. The importance of falls in healthcare will increase substantially with the ageing global population. Falls occurring during hospital admissions are common and costly and can result in injury and increased length of stay (1). In Australia approximately 14% of older people fall during inpatient rehabilitation (2, 3).

Routine use of inpatient fall risk screening tools is suggested in some guidelines (4, 5). The commonly-used Ontario Modified STRATIFY (OMS) (6) is a revised version of the STRATIFY falls screening tool with 5 items (history of falls, mental status, vision, toileting, and transfers/mobility) that are predictive of falls in acute or mixed settings (7). Predict\_FIRST (Prediction of Falls In Rehabilitation Settings Tool) (2) was developed by some of the present authors and contains 5 items (male sex, central nervous system (CNS) medication use, falls in the past year, frequent toileting and inability to tandem stance). A simple question about past falls has been identified as a strong predictor of future falls (8). Experienced clinicians may accurately predict falls risk without using formal tools (9, 10).

The optimal approach to fall prediction in inpatient rehabilitation is not clear. OMS is widely used in rehabilitation internationally and recommended in Australian best practice guidelines (1), but has not been externally validated in rehabilitation (11). Predict\_FIRST was developed for use in rehabilitation, but has not been externally validated. The benefit of using formal fall risk prediction tools rather than clinical judgement or falls history is unknown. This study aimed to address these gaps by comparing several fall prediction methods for use with adults undertaking inpatient rehabilitation. A valid and quick falls screening tool is more likely to be acceptable to clinicians than a longer tool. This study also explored the validity of a shorter fall prediction tool.

The aims of the study were: to test the external validity of 4 approaches to fall prediction in a rehabilitation setting (Predict\_FIRST, Ontario Modified STRATIFY (OMS), physiotherapists' judgement of fall risk (PT\_Risk), and falls in the past year (Past\_Falls); and to develop and test the validity of a simpler fall prediction tool in rehabilitation (Predict\_CM2).

## METHODS

The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines were used in reporting this study (12, 13). Therefore the design of the present study, OMS and Predict\_FIRST development studies were compared (see Supplementary Material<sup>1</sup>).

### Design

Prospective inception cohort study. The study was approved by the SWSLHD ethics committee(s) (HREC number QA2010/005). Approval was granted to audit patient records without gaining written consent from patients.

### Participants

A consecutive sample of 300 inpatients admitted to the general rehabilitation unit at a public hospital in Sydney, Australia participated in the study. Recruitment occurred between April 2010 and May 2011. All admitted patients were considered for inclusion except those who were not receiving rehabilitation, e.g. acute medical or palliative care patients.

### Predictor variables

The lead author (AV) extracted demographic data (primary diagnosis, age and sex) and falls experienced in the previous 12 months from participants' medical records. Medications were transcribed from the medication chart.

The OMS was scored by rehabilitation nursing staff on admission as part of usual care and data for this study were extracted from participants' medical records. The Sydney\_Scoring version of the OMS was used. This version was devised by Lord and colleagues (1) in response to clinician feedback that the transfer and mobility components of the original OMS were confusing to score. The Sydney\_Scoring comprises a simple modification to the mobility score (MS) and transfer score (TS) so that scores on these items are graded in the same direction as all other OMS items. This modification (devised solely for ease of use of the scale) provides the same TS + MS weighted sub-score or overall OMS score as the original scale.

Three versions of the OMS were tested in this study: (i) OMS (Total\_Score) using the continuous total score (/30); (ii) OMS (Dichotomous) using the dichotomous scoring/categories from the original OMS ("At risk" [ $<9$ ] or "No risk" [ $\geq 9$ ]); and (iii) OMS (Risk\_Categories) using low (0–5), medium (6–16) and high (17–30) risk categories (1). OMS (Risk\_Categories) has not been previously validated, so our study also tested the validity of this tool. OMS items were defined in accordance with the original paper for all versions (6).

History of falls was defined as whether the patient presented to hospital with a fall, had fallen since admission or fallen within the last 2 months (6). Impaired mental status was defined as being confused, disorientated or agitated (6). Poor vision was defined as the patient requiring glasses continuously; having blurred vision; or glaucoma, cataracts or macular degeneration (6). Toileting assistance was defined as alterations in urination, i.e. frequency, urgency, incontinence and nocturia (6). Impaired transfers/mobility was defined as needing at least the help of one person to walk and major help of one person to transfer (the use of aids is allowed) (6).

Predict\_FIRST scores were calculated on admission using information from participants' medical records. Male sex was extracted from the file. CNS medication use was defined as taking sedatives/hypnotics, anti-anxiety agents, antipsychotics, antidepressants, anticonvulsants, movement disorder medications or other CNS agents (2). Falls in the past year was defined as a reported or documented history of falls in the previous 12 months (2). Frequent toileting was defined as alterations in urination, i.e. frequency, urgency, incontinence and nocturia (2). Impaired tandem stance was defined as the inability to maintain the tandem stance position for 10 seconds on initial physiotherapy assessment (2).

On admission to rehabilitation, the treating physiotherapist rated and recorded the patient's likelihood of falling during rehabilitation with a simple yes/no response.

### Outcome measures

**Primary outcome.** Number of falls during rehabilitation stay. A fall was defined as unintentionally coming to rest on the ground or other lower surface without overwhelming external force or major internal event (14). Falls were monitored and recorded by the lead author (AV) during the admission from incidents reported in medical records and the ward's fall incidents book, both completed as part of usual care. Supplementary<sup>1</sup> checks of the hospital incident reporting system and regular communication with treating physiotherapists were conducted to optimize falls ascertainment (15). Falls reported by staff and/or recorded in the clinical record, but not entered in the hospital incident reporting system were included in the analysis.

**External validation study.** Discrimination, i.e. the ability of a model to distinguish fallers from non-fallers (16), was assessed using the area under the receiver-operating characteristic (ROC) curve (AUC) and AUCs for different models were compared with the `roccomp` command in Stata 13. The `roccomp` command was used to test the equality of the ROC areas obtained from applying each of the tools to our sample. The command uses an algorithm developed by DeLong and colleagues producing a  $\chi^2$  statistic and associated  $p$ -value (17). A  $p$ -value  $< 0.05$  was interpreted as 2 tools being significantly different in their predictive ability, with the tool with the higher AUC being significantly better at predicting falls. Predicted and observed probability of falling for individuals with different scores on OMS and Predict\_CM2 were calculated.

**Predict\_CM2 development study.** To develop the Predict\_CM2 and determine its ability to discriminate between fallers and non-fallers, we used the same sample as in our external validation study of 300 people.

Associations between individual predictor variables (from Predict\_FIRST and OMS) and falls were assessed using univariate logistic regression analyses from which odds ratios (OR) and 95% confidence intervals (95% CI) were obtained. Variables measuring similar constructs were grouped into domains: past falls; impaired cognition; impaired vision; frequent toileting; impaired mobility; CNS medication use; and sex. Five of these variables from each domain for which the individual  $p$ -values were  $\leq 0.2$  and the odds ratios were greater than 1.5 when dichotomized at the median were identified as candidate predictor variables for multivariate logistic regression models. To increase the confidence in variable selection for the new tool multi-variable analyses were repeated on boot-strapped samples with user written code in Stata 13, to establish the proportion of models in which each predictor was retained. The boot-strapping used 1,000 samples, each with  $n=300$  and involved separate backwards selection

<sup>1</sup><https://doi.org/10.2340/16501977-2290>

logistic regression models in which variables with a  $p$ -value  $< 0.2$  was removed from the models. This was repeated for each sample as the percentage of models in which each variable as retained was reported. Predict\_CM2 was developed selecting measures retained in at least 75% of samples.

*Calibration of existing tools and new tool (Predict\_CM2).* Calibration (the extent to which predicted probabilities agree with observed probabilities) (16) of Predict\_CM2 and OMS was tested with the Hosmer-Lemeshow statistic in Stata 13. A  $p$ -value  $< 0.05$  was interpreted as indicating the model did not fit the data. Multi-level (stratum-specific) likelihood ratios were also calculated.

## RESULTS

### External validation study

*Flow of participants through the study.* During the study 331 patients were admitted to the participating ward. Of these, 300 were eligible for and included in the study (Table I). Thirty-one patients were excluded as they were admitted for acute medical care and did not receive rehabilitation. Complete data were available for all 300 participants.

*Falls.* Forty-one people (14%) experienced 53 falls during their rehabilitation stay. Ten people (3%) had  $\geq 2$  falls and eight people (3%) had  $\geq 3$  falls.

*Individual items and falls.* Univariate associations between Predict\_FIRST and OMS items and faller status are shown in Table II. Variables with the strongest association ( $OR \geq 2.0$ ,  $p < 0.05$ ) with falls included impaired cognition, frequent toileting and impaired mobility/transfers.

*Predictive ability of Predict\_FIRST, OMS, PT\_Risk and Past Falls.* The AUC for Predict\_FIRST was 0.66 (95% CI: 0.57–0.74), lower than the AUC from the tool's development and internal validation study (AUC: 0.73, 95% CI 0.68–0.79). The AUC for OMS

(Total\_Score) was 0.71 (95% CI: 0.61–0.81), higher than, but not statistically different from, Predict\_FIRST ( $p = 0.28$  for comparison). The AUC was not reported in the internal validation study for OMS (6) and therefore cannot be compared.

The AUC for OMS (Total\_Score) was compared with other scoring versions of OMS. Using dichotomous scoring ("At risk" [ $< 9$ ] or "No risk" [ $\geq 9$ ]), OMS (Dichotomous) had an AUC of 0.64 (95% CI: 0.57–0.72), significantly lower than OMS (Total\_Score) ( $p = 0.003$  for comparison). Using the risk category scoring (high, medium, low), OMS (Risk\_Categories) had an AUC of 0.69 (95% CI: 0.60–0.78), lower than, but not statistically different from, OMS (Total\_Score) ( $p = 0.27$  for comparison).

Physiotherapists' judgement of fall risk (PT\_Risk) had an AUC of 0.65 (95% CI: 0.57–0.73), lower than, but not statistically different from, OMS (Total\_Score) ( $p = 0.22$ ). Falls in the last 12 months was a poor predictor of future faller status, with an AUC of 0.52 (95% CI: 0.46–0.60), significantly lower ( $p < 0.001$ ) than OMS (Total\_Score) (see Fig. 1).

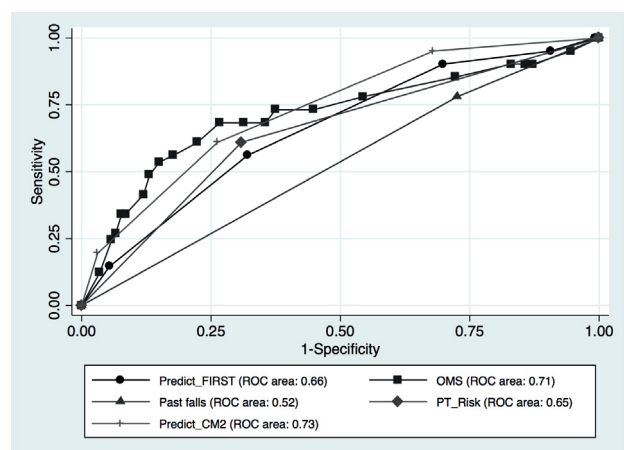
### Predict\_CM2 Development Study

*Development of the predict\_CM2.* The analyses to determine predictors of falling compared 41 fallers with 259 non-fallers, as shown in Table II. The 5 variables that were found to be significant predictors of falls in the univariate analysis (male sex, mobility/transfers, cognition/mental status, CNS medication use, and frequent toileting) were entered into a multivariate regression model. Only these of 3 variables were also found to be significant predictors of falls in the multivariate logistic regression model (male sex, mobility/transfers, and cognition/mental status) and

**Table I.** Demographics of sample ( $n = 300$ , 41 fallers)

	$n$ (%)
Age, years, mean (SD)	80 (11)
Female sex, $n$ (%)	173 (58)
Over 65 years, $n$ (%)	275 (92)
Primary diagnosis, $n$ (%)	
Fracture post-fall	140 (47)
Medical	44 (15)
Fall or syncope	37 (12)
Orthopaedic	22 (7)
Surgical	20 (7)
Respiratory	14 (5)
Neurosurgical	13 (4)
Cardiac	8 (3)
Other	2 (1)
Rehabilitation ward length of stay, days, mean (SD); median (IQR)	22 (13); 20 (15)
Total hospital length of stay, days, mean (SD); median (IQR)	47 (44); 37 (27)

SD: standard deviation; IQR: interquartile range.



**Fig. 1.** Area under receiver-operating characteristic (ROC) Curve (AUC) for Predict\_CM2, Ontario Modified STRATIFY and other predictors ( $n = 300$ , 41 fallers). OMS: Ontario Modified STRATIFY (Total\_Score).

**Table II.** Predict\_FIRST and Ontario Modified STRATIFY items and scores by faller status ( $n = 300$ , 41 fallers)

Predictors	Fallers	Non-fallers	Odds ratio (95% CI)	p-value
	( $n = 41$ ) $n$ (%)	( $n = 259$ ) $n$ (%)		
<b>Predictors</b>				
History of falls <sup>a</sup>	30 (73)	170 (66)	1.34 (0.61–2.95)	0.46
Falls in the past year <sup>b</sup>	32 (78)	188 (73)	1.43 (0.68–2.98)	0.34
Impaired cognition <sup>a</sup>	23 (56)	56 (22)	4.63 (2.34–9.18)	<0.001
Impaired vision <sup>a</sup>	17 (41)	120 (46)	0.80 (0.42–1.52)	0.49
Frequent toileting	24 (59)	104 (40)	2.10 (1.08–4.11)	0.03
Impaired mobility/ transfers <sup>a</sup>	26 (63)	92 (36)	3.15 (1.59–6.24)	<0.001
CNS medication <sup>b</sup>	30 (73)	161 (62)	1.66 (0.80–3.46)	0.18
Unable to do tandem stance <sup>b</sup>	37 (90)	213 (82)	2.00 (0.68–5.88)	0.21
Male sex <sup>b</sup>	23 (56)	104 (40)	1.90 (0.98–3.70)	0.06
<b>Predict_CM2 score (/3)</b>				
0	2 (5)	83 (32)		
1	14 (34)	108 (42)		
2	17 (41)	60 (23)		
3	8 (20)	8 (3)		
<b>Ontario Modified STRATIFY score</b>				
No risk (<9)	11 (27)	143 (55)		
At risk (≥9)	30 (73)	116 (45)		
Mean (SD)	10.7 (8)	18.2 (10)		

Odds ratios (OR) are for the odds of having at least one fall from logistic regression models.

<sup>a</sup>Item from Ontario Modified STRATIFY only. <sup>b</sup>Item from Predict\_FIRST only. CI: confidence interval; CNS: central nervous system.

*History of falls:* Defined as whether the patient presented to hospital with a fall, fallen since admission or fallen within the last 2 months (6). *Falls in the past year:* Defined as whether there is a reported or documented history of falls in the previous 12 months (2). *Impaired cognition:* Defined as being confused, disorientated or agitated (6). *Impaired vision:* Defined as the patient requiring glasses continuously; having blurred vision; or glaucoma, cataracts or macular degeneration (6). *Frequent toileting:* Defined as alterations in urination, such as frequency, urgency, incontinence and nocturia (6). *Impaired mobility:* Defined as needing at least the help of 1 person to walk and major help of 1 person to transfer (the use of aids is allowed) (6). *CNS medication:* Defined as taking sedatives/hypnotics, anti-anxiety agents, antipsychotic agents, antidepressants, anticonvulsants, movement disorder medications or other CNS agents (2). *Unable to do tandem stance:* Defined as the inability to maintain the tandem stance position for 10 s on initial physiotherapy assessment (2).

**Table III.** Odds ratios (OR) and coefficients from the Predict\_CM2 multivariate models ( $n = 300$ , 41 fallers). The table shows ORs and adjusted and zero-adjusted coefficients from 1,000 boot-strapped samples

Predictor	% of models on boot-strapped samples in which retained	Zero-adjusted coefficients from multivariable analysis on boot-strapped samples <sup>c</sup>	Odds ratios (95% CIs) from multivariable analysis on the original sample
Impaired cognition <sup>a</sup>	99.0	1.4	4.16 (2.05–8.42)
Impaired mobility <sup>b</sup>	88.5	0.9	2.74 (1.34–5.62)
Male sex	78.2	0.7	2.11 (1.04–4.29)

<sup>a</sup>Defined as being confused, disorientated or agitated (6). <sup>b</sup>Defined as needing at least the help of one person to walk and major help of one person to transfer (the use of aids is allowed) (6). <sup>c</sup>Mean coefficients from the multivariate models in all of the boot-strapped samples with zero allocated for samples in which the variable was not retained. Note: Frequent toileting was retained in 38.1% and CNS medication in 42.2% of boot-strapped samples.

**Table IV.** Predicted probability of falling and numbers of people actually falling for each Predict\_CM2, Ontario Modified STRATIFY (Risk\_Categories), Ontario Modified STRATIFY (Dichotomous) and multilevel (stratum-specific) likelihood ratios for each level of test ( $n = 300$ , 41 fallers)

Tool	People with this score, $n$	Predicted probability of falling, %	Actual probability of falling, %	People who actually fell, $n$	Likelihood ratio (95% CI)
<b>Predict_CM2 score (0–3)</b>					
Very low risk (0)	85	3	2	2	0.15 (0.04–0.51)
Low risk (1)	122	10	11	14	0.82 (0.51–1.23)
Moderate risk (2)	77	24	22	17	1.79 (1.14–2.74)
High risk (3)	16	48	50	8	6.32 (2.55–15.24)
<b>OMS (Risk_Categories)</b>					
Low (0–5)	48	3	8	4	0.57 (0.22–1.39)
Medium (6–16)	183	10	8	14	0.52 (0.33–0.77)
High (17–30)	69	30	33	23	3.16 (2.12–4.52)
<b>OMS (Dichotomous)</b>					
No risk (<9)	154	7	7	11	0.49 (0.28–0.77)
At risk (≥9)	146	21	21	30	1.63 (1.26–2.01)

were retained in more than 75% of multi-variate models tested on 1,000 boot-strapped samples (Table III). The final tool was created using these 3 variables and named Predict\_CM2 based on the first letter of each of the included variables (i.e. cognition, male, mobility).

*Predictive ability of the Predict\_CM2.* The simple tool with risk factors equally (unit) weighted had an AUC of 0.73 (95% CI: 0.66–0.81, coefficient from logistic regression model 1.08, constant –3.31). The AUC for Predict\_CM2 was significantly higher than Predict\_FIRST (AUC: 0.66, 95% CI: 0.57–0.74,  $p = 0.04$  for comparison) or Past\_Falls (AUC: 0.52, 95% CI 0.46–0.60,  $p < 0.01$ ) and higher than, but not statistically different from, OMS (Total\_Score) (AUC: 0.71, 95% CI: 0.61–0.81,  $p = 0.52$  for comparison) and PT\_Risk (AUC: 0.65, 95% CI: 0.57–0.73,  $p = 0.08$  for comparison). The AUC for Predict\_CM2 was significantly higher than OMS (Dichotomous) (AUC: 0.64, 95% CI: 0.57–0.72,  $p = 0.008$  for comparison) and higher than, but not statistically different from, OMS (Risk\_Categories) (AUC: 0.69, 95% CI: 0.60–0.78,  $p = 0.17$  for comparison).

*Absolute probability of falling and tool calibration (Predict\_CM2 and OMS)*

Table IV shows the absolute probability of falling, with different scores on the different tools. A person with no Predict\_CM2 risk factors (0/3) had a 3%, or very low probability of falling in hospital, while those with 3 risk factors had a 48%, or high probability. The

resulting tool is shown in Appendix I. A person in the OMS (Dichotomous) “No risk” category (0–8/30) had a 7% probability of falling in hospital, while those in the “At risk” category (9–30/30) had a 21% probability. A person in the OMS (Risk\_Categories) “Low risk” category (0–5/30) had a 3% probability of falling in hospital, while those in the “High risk” category (17–30/30) had a 30% probability.

The Hosmer-Lemeshow test did not detect a lack of fit between predicted and observed fallers for Predict\_CM2 or OMS (Predict\_CM2:  $p=0.62$ , OMS Dichotomous:  $p<0.01$  and OMS Risk\_Categories:  $p=0.57$ ). Likelihood ratios (Table IV) for Predict\_CM2 (0.15–6.32) have a greater range than for OMS (Total\_Score) (0.49–1.63) suggesting a greater usefulness of this tool.

## DISCUSSION

This study found relatively poor performance of Predict\_FIRST (AUC=0.66), PT\_Risk (AUC=0.65) and Past\_Falls (AUC=0.52) at predicting falls in inpatient rehabilitation. While fall history has been found to be a predictor of falls in similar populations and settings, our study, as well as a more recent study, has found it to be a poor predictor of falls in rehabilitation (18). The performance of the OMS (Total\_Score) (AUC=0.71) provides good external validation of this tool and demonstrates it is significantly better in predicting falls than “Past\_Falls” ( $p<0.001$ ). Given the time taken to administer the 5-item (with 11 sub-questions) OMS and relatively poor performance of the other screens, we developed a simpler 3-item tool for predicting falls in rehabilitation, Predict\_CM2, on the basis of the strongest individual predictors in this sample. The predictive ability of Predict\_CM2 (AUC=0.73) was comparable to OMS (Total\_Score), approached significance when compared with PT\_Risk ( $p=0.09$ ) and was significantly better than Predict\_FIRST ( $p=0.04$ ), or Past\_Falls ( $p<0.001$ ) in predicting future fallers.

This study used data routinely collected from participants, whereas the Predict\_FIRST development study involved additional assessment requiring consent. These design differences appear to have led to important differences in the results, with the present sample more representative of routine care. Only 3 predictor variables used in Predict\_FIRST and OMS were strongly associated with falls in the present sample: mobility/transfers, mental status/cognition, and male sex. Cognition was not a strong predictor of falls in the Predict\_FIRST study, probably due to exclusion of people with cognitive impairment for whom no person responsible was available to give consent. Tandem stance inability and CNS medication use were

not strong predictors of falls in this sample, probably due to a higher proportion of participants with these risk factors compared with the Predict\_FIRST study. We therefore propose a simplified fall prediction tool (Predict\_CM2) including the 3 variables, mobility/transfers, mental status/cognition, and male sex, which are also predictive of falls in similar settings (8, 18–20).

The predictive ability of Predict\_CM2 or OMS (Total Score) was better, but not statistically different (AUC=0.73 and 0.71, respectively) from clinical judgement alone (AUC=0.65). Previous studies have found clinical judgement to be a comparable or better predictor of falls than formal assessment tools (9, 10, 21). We suggest these differences may be statistically significant in a larger sample and by using a screening tool, clinicians with differing backgrounds and levels of experience are likely to more confidently and accurately determine the patient’s level of risk. The ability of Predict\_CM2 and now OMS (Dichotomous and Risk Category versions) to provide an absolute risk of falling as a percentage may assist in communicating level of risk to staff as well as to patients and significant others.

The use of formal fall risk screening tools is not supported by all (4). UK guidelines recommend a multifactorial assessment, and do not support using risk prediction tools (5). American and British guidelines recommend fall risk screening by asking about past falls, fall frequency and difficulties in gait and balance (22). We suggest using a simple externally-validated tool addresses some concerns raised about using fall prediction tools. First, the Predict\_CM2 provides an absolute risk of falling, which gives the clinician more information about falls risk than tools such as the OMS that provide risk categories alone. Secondly, using the tool will identify more patients at risk of falling than clinical judgement or past falls. Thirdly, such a tool can also be used to guide provision of interventions, but this approach should be evaluated in trials. For example, patients at high risk with cognitive impairment may need closer monitoring, relocating closer to the nurses’ station or require a special nurse. Patients with impairments in mobility and transfers may require more assistance when carrying out activities of daily living, and rehabilitation can be targeted to improve mobility and transfers.

Guidelines recommend that independent investigators should externally validate prediction models (23). We have independently validated the OMS, and while the investigators who developed Predict\_FIRST were involved in external validation, 2 independent investigators were also involved. We recognize in developing a new tool, external validation of Predict\_CM2 is required (24). Any comparisons with other predictors are likely to favour the Predict\_CM2 and should be

interpreted with caution. We recognize that our setting had a relatively high proportion of patients with fractures/orthopaedic conditions and a low proportion of patients with neurological conditions compared with other rehabilitation settings. Further validation of the tool is recommended in rehabilitation settings with a higher neurological case mix.

As we were unable to collect data regarding the number of past falls or fall-related injuries, we could not explore comparisons between single and multiple fallers or injurious and non-injurious falls. Other limitations include the relatively small sample size and use of a single site. Some suggest that external validation studies should involve a minimum of 100 events and 100 non-events (25). A sample of this size was not feasible for the present study, given the limited resources available.

The use of multi-modal prospective approaches to capturing falls in hospitals is recommended due to the generally poorer accuracy of single or retrospective approaches (26, 27). The use of multi-modal prospective methods to report falls is a strength of the present study. The systematic approach used to select items for the tool overcomes the risk of spurious findings in small samples; inclusion of participants with cognitive impairment improves generalizability; and the availability of complete data for all 300 participants minimizes the risk of bias associated with incomplete follow-up.

In conclusion, Predict\_CM2 was significantly better than Predict\_FIRST and a single question on past falls in predicting falls in rehabilitation. Predict\_CM2 was better, but not statistically different from OMS (Total\_Score) or PT\_Risk at predicting falls in rehabilitation. Predict\_CM2 is a simplified screening tool, providing a quick way to quantify the probability of a patient falling during rehabilitation. After external validation, Predict\_CM2 could be used routinely in an inpatient rehabilitation setting.

### ACKNOWLEDGEMENTS

This study was supported by an infrastructure grant from the Ingham Institute for Applied Medical Research. AT, SL and CS are supported by Fellowships from the Australian National Health and Medical Research Council.

### REFERENCES

1. Australian Commission on Safety and Quality in Healthcare. Preventing falls and harm from falls in older people: best practice guidelines for Australian hospitals. Sydney, Australia: Australian Commission on Safety and Quality in Healthcare; 2009.
2. Sherrington C, Lord SR, Close JCT, Barraclough E, Taylor M, Oa Rourke S, et al. Development of a tool for prediction

- of falls in rehabilitation settings (Predict\_FIRST): a prospective cohort study. *J Rehabil Med* 2010; 42: 482–488.
3. Cumming RG, Sherrington C, Lord SR, Simpson JM, Vogler C, Cameron ID, et al. Cluster randomised trial of a targeted multifactorial intervention to prevent falls among older people in hospital. *BMJ* 2008; 336: 758–760.
4. Oliver D. Falls risk-prediction tools for hospital inpatients. Time to put them to bed? *Age Ageing* 2008; 37: 248–250.
5. National Institute for Health and Care Excellence. Falls: assessment and prevention of falls in older people. Manchester, UK: National Institute for Health and Care Excellence; 2013.
6. 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.
7. 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.
8. Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ. Will my patient fall? *JAMA* 2007; 297: 77–86.
9. Lundin-Olsson L, Jensen J, Nyberg L, Gustafson Y. Predicting falls in residential care by a risk assessment tool, staff judgement, and history of falls. *Aging Clin Exper Res* 2003; 15: 51–59.
10. 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: 277–281.
11. Billington J, Fahey T, Galvin R. Diagnostic accuracy of the STRATIFY clinical prediction rule for falls: a systematic review and meta-analysis. *BMC Fam Pract* 2012; 13: 76.
12. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med* 2015; 13: 1.
13. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; 162: W1–W73.
14. The prevention of falls in later life. A report of the Kellogg International Work Group on the Prevention of falls by the elderly. *Danish Med Bull* 1987; 34: 1–24.
15. Hill A-M, Hoffmann T, Hill K, Oliver D, Beer C, McPhail S, et al. measuring falls events in acute hospitals—a comparison of three reporting methods to identify missing data in the hospital reporting system. *J Am Geriatr Soc* 2010; 58: 1347–1352.
16. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statist Med* 1996; 15: 361–387.
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44: 837–845.
18. Aizen E, Zlotver E. Prediction of falls in rehabilitation and acute care geriatric setting. *J Clin Gerontol Geriatr* 2013; 4: 57–61.
19. Vieira ER, Freund-Heritage R, da Costa BR. Risk factors for geriatric patient falls in rehabilitation hospital settings: a systematic review. *Clin Rehabil* 2011; 25: 788–799.
20. Healey F, Scobie S, Oliver D, Pryce A, Thomson R, Glampson B. Falls in English and Welsh hospitals: a national observational study based on retrospective analysis of 12 months of patient safety incident reports. *Qual Safety Health Care* 2008; 17: 424–430.
21. Haines TP, Hill K, Walsh W, Osborne R. Design-related bias in hospital fall risk screening tool predictive accuracy evaluations: systematic review and meta-analysis. *J Gerontol*

- Ser A-Biol Sci Med Sci 2007; 62: 664–672.
22. Panel on Prevention of Falls in Older Persons AGS, British Geriatrics S. Summary of the updated American Geriatrics Society/British Geriatrics Society Clinical Practice Guideline for Prevention of Falls in Older Persons. *J Amer Geriatr Soc* 2011; 59: 148–157.
  23. Collins G, de Groot J, Dutton S, Omar O, Shanyinde M, Tajar A, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol* 2014; 14: 40.
  24. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ* 2009; 338: b375.
  25. Vergouwe Y, Steyerberg EW, Eijkemans MJC, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005; 58: 475–483.
  26. Hauer K, Lamb SE, Jorstad EC, Todd C, Becker C. Systematic review of definitions and methods of measuring falls in randomised controlled fall prevention trials. *Age Ageing* 2006; 35: 5–10.
  27. Hill AM, Hoffmann T, Hill K, Oliver D, Beer C, McPhail S, et

al. Measuring falls events in acute hospitals – a comparison of three reporting methods to identify missing data in the hospital reporting system. *J Am Geriatr Soc* 2010; 58: 1347–1352.

#### Appendix I. The Predict\_CM2 Tool

	Score
Impaired mobility/transfers <sup>a</sup>	1
Impaired cognition <sup>b</sup>	1
Male	1
Total score	/3
Probability of falling with different scores:	
0 = 3%	Very low risk
1 = 10%	Low risk
2 = 24%	Moderate risk
3 = 48%	High risk

<sup>a</sup>Defined as needing at least the help of 1 person to walk and major help of 1 person to transfer (the use of aids is allowed) (6).

<sup>b</sup>Defined as being confused, disorientated or agitated (6).