Factors associated with fall risk of community-dwelling older people: A decision tree analysis

Kenneth N K Fong[1] 🆔 **, Raymond C K Chung[1], Patrick P C Sze[2] and Carmen K M NG[2]**

Abstract

Objective: To examine the predictive attributes for accidental falls in community-dwelling older people in Hong Kong using decision tree analysis.

Methods: We recruited 1151 participants with an average age of 74.8 years by convenience sampling from a primary health-care setting to carry out the cross-sectional study over 6 months. The whole dataset was divided into two sets, namely training set and test set, which respectively occupied 70% and 30% of the whole dataset. The training dataset was used first; decision tree analysis was used to identify possible stratifying variables that could help to generate separate decision models.

Results: The number of fallers was 230 with 20% 1-year prevalence. There were significant differences in gender, use of walking aids, presence of chronic diseases, and co-morbidities including osteoporosis, depression, and previous upper limb fractures, and performance in the Timed Up and Go test and the Functional Reach test among the baselines between the faller and non-faller groups. Three decision tree models for the dependent dichotomous variables (fallers, indoor fallers, and outdoor fallers) were generated, with overall accuracy rates of the models of 77.40%, 89.44% and 85.76%, respectively. Timed Up and Go, Functional Reach, body mass index, high blood pressure, osteoporosis, and number of drugs taken were identified as stratifying variables in the decision tree models for fall screening.

Conclusion: The use of decision tree analysis for clinical algorithms for accidental falls in community-dwelling older people creates patterns for decision-making in fall screening, which also paves the way for utility-based decision-making using supervised machine learning in fall risk detection.

Keywords

Accidental falls, older people, community-dwelling, decision tree analysis, machine learning

Submission date: 11 September 2022; Acceptance date: 24 May 2023

Introduction

The population in Hong Kong is aging rapidly as a result of its having the highest life expectancy in the world. In 2021, the life expectancy at birth of males was 83.2 years while that of female was 87.9 years.[1] Hong Kong is also the world's most densely populated city, packed with high-rise buildings, with half of its people living above the 15th floor in multi-story residential buildings[2] and nearly half living in

[1]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong SAR
[2]Elderly Resources Centre, Hong Kong Housing Society, Hong Kong SAR

Corresponding author:
Kenneth N K Fong, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong SAR.
Email: rsnkfong@polyu.edu.hk

public housing.[3] Falls are common among the older population resided in community dwellings in Hong Kong – the fall rate has been reported as 29% annually, and the once-yearly prevalence of falls has been reported as 20%, dropping to 6.3% for 2 or more falls per year. Of all falls, 47.7% occurred indoors while 52.3% occurred outdoors.[4] The figures are less than that of 29% occur at least once annually reported in the United States.[5]

The causes of falls are multifactorial and result from a combination of intrinsic and extrinsic factors; among intrinsic factors, impairments in gait and balance are the most prominent predisposing risk factors at the population level, followed by visual impairment, orthostatic hypotension, and polymedication; on the other hand, osteoporosis is an important contributing cause of fall-related fractures.[6] According to the World guidelines for falls prevention, there is a need for healthcare providers to develop a standard approach to assess an individual's estimated level of risk for falls for fall risk stratification in order to apply an appropriate intervention according to level of risk, particularly an objective assessment of gait and balance for differentiating immediate and high from low risk of falls is strongly recommended.[7] Such tools have been shown to reduce falls and fall-related injuries for community-dwelling older people.[8,9] The use of a minimum number of measures in screening would positively affect outcomes but also ensure a minimal effect on staff time required or the need for additional equipment.[10] Decision tree model had been used to convert a complex decision-making process based on varying influences of significant risk factors for falls in large cohorts of older adults, into a fall-risk classification algorithm that is followed in a stepwise manner to reach a final conclusion for clinical application and interpretation.[11,12] This cross-sectional study aimed to examine the predictive attributes based on a fast screening assessment comprised of 60 variables, comprising medical, social and physical outcomes, including functional assessments for gait and balance, for accidental falls in community-dwelling older people in Hong Kong using decision tree analysis.

Methods

Participants

This was a cross-sectional study for two prospective screening occasions took place in 2009–2010 and 2014–2015. Older people aged 65 years or above attending a primary healthcare setting – the elderly resource centre (the Centre) of the Hong Kong Housing Society – were recruited for assessment. This is a primary healthcare centre aiming to provide prevention in the form of education to older people residing in community dwellings who are at risk. The participants in this study were clients referred from different community organizations and visiting the Centre

once from various districts in Hong Kong during the study period. Baseline assessments were completed at the Centre with informed and written consent from each participant before data collection. Inclusion criteria for this study were: (a) living independently in community dwellings and ambulatory with or without a walking aid, and (b) having a Mini-Mental State Examination score ≥ 18.[13] Those who were unable or unwilling to give consent were excluded from the study, as were those residing in nursing homes. Those with poorer cognitive performance were excluded because they would have been less reliable in their recall and ability to report their fall history.[14] To obtain a more representative sample, the study population was divided into five strata, defined by age according to the distribution of the Hong Kong Population Census of 2012.[15] The study was approved by the Human Subjects Ethics Committee of the Hong Kong Polytechnic University (Ref: HSEARS20140628001) and the governing board of the Centre.

In this study, a fall was defined as an event that results in a person coming to rest inadvertently on the ground or hitting an object like a chair or stair.[16] Participants who fell because of excess alcohol intake, or who fell after sustaining a sudden blow or loss of consciousness or sudden onset of paralysis due to stroke or epileptic seizure, were excluded. If the participants reported a fall in the preceding 12 months, they would be further asked about the frequency of falls, injuries from falls, and indoor or outdoor falls. Indoors was operationally defined as the flat in which the person was living and the environment within the house, apartment, or building, but excluding the common areas outside the building. Apart from age and gender, level of education, previous injury, comorbidity (please refer to Table 1), performance of instrumental activities of daily living (IADL) – ability to perform light and heavy household duties, number of drugs taken, activity level, mobility status, use of walking aid, and social support were self-reported and recorded. The reason for using self-reporting was that the participants attended the Centre just once and that medical information would not be accessible in a community centre. Other assessments included measurements of body mass index (BMI) – calculated as weight (kg) divided by the square of height (m) – and blood pressure (BP), the Timed Up and Go (TUG) test,[17] the Functional Reach (FR) test,[18] and a visual acuity test using a Snellen Eye Chart.[19] Both TUG and FR had been used for risk assessment screening in a similar large-scale nation-wide survey of the older population.[20] The TUG is an objective and reliable gait measurement that measures the time taken by an individual to stand up from a standard armchair, walk for 3 m, turn around and walk back to the chair and sit down again.[21] The recent guideline recommend a cut-off value of >15 s for the TUG for older people with gait and balance impairment.[7] The FR is a reliable measure of balance for older community-dwelling populations and

Table 1. Baseline demographics of the population and comparison between fallers and non-fallers.

Characteristics	All (*n* = 1151)	Fallers (*n* = 230)	Non-fallers (*n* = 921)	*p* value
Gender, n (%)	1151 (100.0)	230 (100.0)	921 (100)	0.000**
Male	231 (20.1)	25 (10.9)	206 (22.4)	
Female	920 (79.9)	205 (89.1)	715 (77.6)	
Age, mean ± SD	74.8 ± 6.71	75.06 ± 6.7	74.75 ± 6.7	0.525
Age range, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	
65–69	306 (26.6)	55 (23.9)	251 (27.3)	
70–74	299 (26.0)	60 (26.1)	238 (25.8)	
75–79	257 (22.3)	57 (24.8)	201 (21.8)	
80–84	170 (14.8)	31 (13.5)	139 (15.1)	
85 +	119 (10.3)	27 (11.7)	92 (10.0)	
Education, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.596
No formal education	414 (35.9)	88 (38.3)	326 (35.4)	
Primary	399 (34.7)	83 (36.1)	316 (34.3)	
Lower secondary	109 (9.5)	18 (7.8)	91 (9.9)	
Upper secondary	114 (9.9)	16 (7.0)	98 (10.6)	
Finish middle school	71 (6.2)	12 (5.2)	59 (6.4)	
Tertiary	44 (3.8)	13 (5.7)	31 (3.4)	
Living environment, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.488
Public housing estate	479 (41.6)	106 (46.1)	373 (40.5)	
Housing ownership scheme	149 (12.9)	29 (12.6)	120 (13.0)	
Private housing	325 (28.2)	58 (25.2)	267 (29.0)	
Rented room	15 (1.3)	1 (0.4)	14 (1.5)	
Senior citizen hostel	75 (6.5)	13 (5.7)	62 (6.7)	
Squatter hut/ temporary house	0 (0)	0 (0)	0 (0)	
Others	108 (9.4)	23 (10)	85 (9.2)	
Living alone, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.226
Yes	357 (31.0)	79 (34.3)	279 (30.3)	

(continued)

Table 1. Continued.

Characteristics	All (n = 1151)	Fallers (n = 230)	Non-fallers (n = 921)	p value
No	794 (69.0)	151 (65.7)	642 (69.7)	
Living with couple only, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.311
Yes	242 (21.0)	54 (23.5)	188 (20.4)	
No	909 (79.0)	176 (76.5)	732 (79.5)	
Dependent for light household duties, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.135
Yes	326 (28.3)	56 (24.3)	270 (29.3)	
No	825 (71.7)	174 (75.7)	651 (70.7)	
Dependent for heavy household duties, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.141
Yes	643 (55.9)	138 (60.0)	503 (54.6)	
No	508 (44.1)	92 (40.0)	418 (45.4)	
Use walking aid, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.000**
Yes	738 (64.1)	78 (33.9)	199 (21.6)	
No	413 (35.9)	152 (66.1)	722 (78.4)	
Presence of chronic diseases, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	0.005**
Yes	902 (78.4)	196 (85.2)	706 (76.7)	
No	249 (21.6)	34 (14.8)	215 (23.3)	
Co-morbidities, n (%)		230 (100.0)	921 (100.0)	
Stroke	45 (3.9)	11 (4.8)	34 (3.7)	0.445
Dementia	1 (0.1)	0 (0)	1 (0.1)	0.617
Osteoporosis	60 (5.2)	19 (8.3)	41 (4.5)	0.020*
Arthritis	100 (8.7)	27 (11.7)	73 (7.9)	0.066
Parkinson's disease	3 (0.3)	0 (0)	3 (0.3)	0.386
High blood pressure	568 (49.3)	124 (53.9)	444 (48.2)	0.769
Diabetes mellitus	218 (18.9)	42 (18.3)	176 (19.1)	
Eye disease	207 (18.0)	52 (22.6)	155 (16.8)	0.041
Low blood pressure	14 (1.2)	3 (1.3)	11 (1.2)	0.892
Chronic chest disease	13 (1.1)	2 (0.9)	11 (1.2)	0.677
Asthma	21 (1.8)	4 (1.7)	17 (1.8)	0.914

(continued)

Table 1. Continued.

Characteristics	All (n = 1151)	Fallers (n = 230)	Non-fallers (n = 921)	p value
Cardiac disease	123 (10.7)	29 (12.6)	94 (10.2)	0.291
Depression	19 (1.7)	8 (3.5)	11 (1.2)	0.015*
Cancer	16 (1.4)	1 (0.4)	15 (1.6)	0.167
Previous upper limb fracture	8 (0.7)	5 (2.2)	3 (0.3)	0.003**
Previous lower limb fracture	13 (1.1)	4 (1.7)	9 (1.0)	0.328
Low back pain	14 (1.2)	6 (2.6)	8 (0.9)	0.031
Others	351 (30.5)	83 (36.1)	268 (29.1)	0.039
Number of drugs intake, n (%)	1151 (100.0)	230 (100.0)	921 (100.0)	
Nil	349 (30.3)	55 (23.9)	294 (31.9)	
1–3	738 (64.1)	154 (67.1)	584 (63.4)	
≥4	64 (5.6)	21 (9.1)	43 (4.7)	
Time Up and Go Test, mean ± SD	14.1 ± 4.3	15.15 ± 5.13	13.78 ± 4.07	0.000**
Forward Reaching (cm), mean ± SD	24.5 ± 6.0	23.61 ± 6.23	24.69 ± 5.94	0.015*
BMI, mean ± SD	24.3 ± 3.7	24.52 ± 3.99	24.21 ± 3.59	0.262
AMT, mean ± SD	8.8 ± 1.5	8.71 ± 1.45	8.76 ± 1.48	0.595
Visual acuity (left), mean ± SD	5.5 ± 1.8	5.48 ± 1.74	5.44 ± 1.80	0.806
Visual acuity (right), mean ± SD	5.6 ± 1.7	5.69 ± 1.53	5.55 ± 1.80	0.271

BMI: body mass index; AMT: Abbreviated Mental Test.
* $p \leq 0.05$; ** $p \leq 0.01$

measuring the difference between arm's length and maximal forward reach using a fixed base of support.[22] The whole assessment process for each participant took 30 min to complete. Five trained occupational therapy students together with two experienced occupational therapists were responsible for carrying out all assessments of the participants.

Statistical analysis

Descriptive analysis was conducted on all baseline characteristics for the sample and the two subgroups, which were fallers and non-fallers according to the 12-months prevalence of falls (persons with at least one fall over the previous 12 months). Comparison of baseline characteristics between fallers and non-fallers was performed using independent-samples *t*-test if the outcome measures were continuous data. Chi-squared test was used instead if the outcome measures were categorical. Level of significance was set at 0.05.

Decision tree analysis was conducted using the rpart package of R software,[23] which was a powerful machine learning (ML) library in R that was used for building both classification and regression trees. This library implemented recursive partitioning and was very easy to use. In our study, our decision tree was a standard classification tree with the predicted values indicating the class to classify the subjects into fallers and non-fallers through a random split of the data into two portions: training and testing data. There were two major concerns to decide the optimum split of the training and testing data. (a) If there was less training data, the ML method would show high variance training; (b) with less testing data, the model evaluation performance statistics would have greater variance. Thus, the whole dataset was divided into two sets, namely training set and test set, which occupied 70% and

30% of the whole dataset, respectively, this was one of the standard and conservative splitting proportions under ML method (80% vs 20%, 70% vs 30% and 60% vs 40%).[24] The training set was used with decision tree analytical software R to identify all possible stratifying variables such as TUG, the FR, the BMI, high BP, osteoporosis, and number of drugs taken that could help to generate separate decision models for each of the three dependent dichotomous variables – fallers, indoor fallers, and outdoor fallers among community-dwelling older adults. The test set was then used to test the validity of these three decision models and their accuracy rates were then computed accordingly. The accuracy rate of the decision tree was computed by the sum of true-positives (fallers) and true-negatives (non-fallers) among the total number of samples within confusion matrix of the testing dataset. Cross-validation was used to evaluate the model with the same train-test split (70% vs 30%) for 10 training iterations. Among the 10 training iterations, the decision tree with the interpretable framework together with at least the median accuracy rate of the 10 decision trees had been chosen for the final decision tree model. Though the hyperparameter tuning such as split criteria, mini samples split, mini samples leaf or maximal depth, etc. through grid search method could help to optimize the decision tree model architecture, it was not a compulsory process but only an optional process for decision tree analysis. Also, it was only a way of improving the decision tree model and thus hyperparameter tuning was not conducted under our decision tree analysis. However, the default settings of the hyperparameters under the rpart library were used in our study.

Apart from adopting the Decision tree analysis, Random Forest model will be used to produce a reasonable prediction without hyperparameter tuning. The performance metric of the model such as accuracy, sensitivity, specificity, positive predictive value and negative predictive value will be computed. As there were many more non-fallers than fallers leading to imbalanced classification in our study, the sampling methods such as undersampling, oversampling, both undersampling and oversampling as well as ROSE function[25] would be applied to our random forest analysis to further enhance its performance metric mentioned above.

Results

Decision tree analysis

A total of 1151 participants, with an average age of 74.8 (SD = 6.7) years, were recruited by convenience sampling to this study over 6 months. The 1-year prevalence of falls was 20% ($n = 230/1151$) and the fall rate (number of falls per 100 persons in one year) was 30.4% ($n = 350/1151$). The mean number of falls per faller was 1.4. The numbers of indoor and outdoor fallers were 96 and 157,

respectively, with 23 participants having both indoor and outdoor falls in 1 year. Among the fallers, 78 of them had recurrent falls (with 2 or more falls) within a year. The 1-year prevalence of recurrent falls was 6.8% ($n = 78/1151$). Table 1 shows the baseline characteristics of the whole sample and the faller and non-faller groups. Among all the baselines, only 8 of them were shown to be significant between the faller and non-faller groups. Regarding gender, females were more likely than males to be fallers (89.1% vs 10.9%, $p < 0.001$). It was also found that those who did not use walking aids tended to be fallers more than those who did use them (66.1% vs 33.9%, $p < 0.001$). Moreover, those who showed manifestations of chronic diseases were more likely to be fallers than those who did not (85.2% vs 14.8%, $p = 0.005$). Fallers were also more likely than non-fallers to have co-morbidity problems like osteoporosis (8.3% vs 4.5%, $p = 0.02$), depression (3.5% vs 1.2%, $p = 0.015$), and previous upper limb fracture (2.2% vs 0.3%, $p = 0.003$. In terms of functional tests like TUG and FR, fallers took longer to complete the former (15.15 vs 13.78, $p < 0.001$) and could not reach as far in the latter (23.61 vs 24.69 cm, p = 0.015) compared with non-fallers. Out of 1151 participants, about 70% of them ($n = 810$) were used for the training dataset and the remainder were used for the test dataset ($n = 341$). Figures 1 to 3 show the decision tree diagrams for the dependent dichotomous variables (fallers, indoor fallers, and outdoor fallers), with overall model accuracy rates of 77.40%, 89.44%, and 85.76%, respectively.

Regarding the decision tree for fallers (Figure 1), 151 out of 797 participants were classified as fallers. Five stratifying variables were identified, namely TUG, FR, BMI, high BP, and osteoporosis. TUG was chosen as the first stratifying variable and the threshold was set at 15. Participants with TUG greater than or equal to 15 were found to have a higher fall rate of 26.67% compared with those with TUG less than 15, who had a relatively lower fall rate of 14.99%. Among those participants with higher TUG, FR was chosen as the second stratifying variable and the threshold was set at 16. Those with FR less than 16 had a higher fall rate of 45.45% compared with those who had a higher FR (greater than or equal to 16), who had a lower fall rate of 24.05%. Among those with FR less than 16, BMI was taken as another stratifying variable, with 22 taken as the cutoff. Those with BMI greater than or equal to 22 were found to have a higher fall rate of 56% compared with those who had BMI less than 22, who had a lower fall rate of 12.5%. Among those participants with BMI greater than or equal to 22, those who had high BP were more likely to become fallers than those without high BP (68.75% vs 33.33%). Among those participants with TUG less than 15, osteoporosis is chosen as the only stratifying variable. Those with osteoporosis had a higher fall rate of 39.13% compared with those without, who had a lower fall rate of 13.89%.
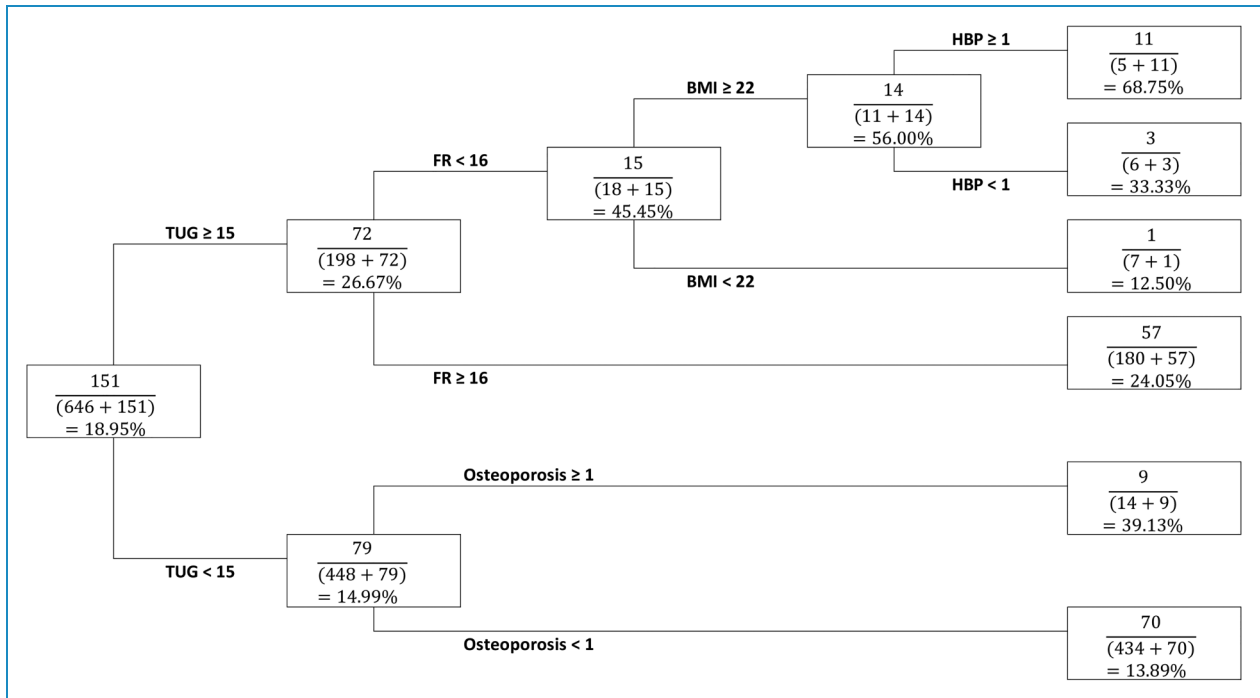
Figure 1. Decision tree for fallers
*overall accuracy rate = 77.40%.
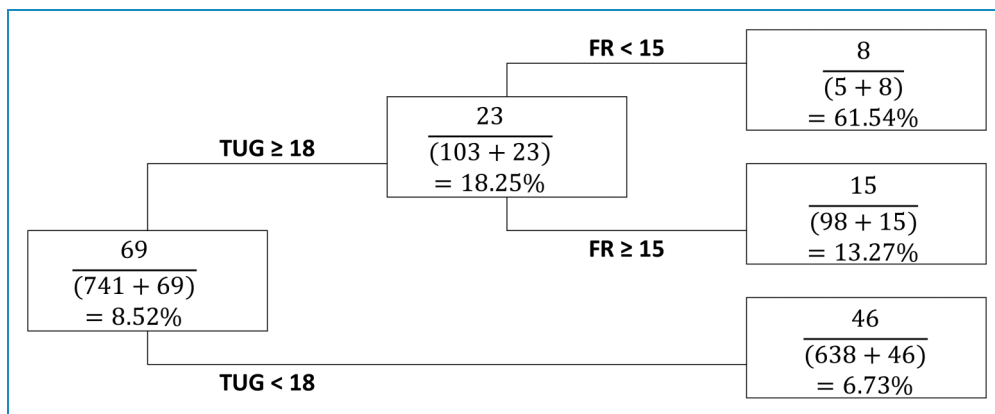
Figure 2. Decision tree for indoor fallers
*overall accuracy rate = 89.44%.

Regarding the decision tree for indoor fallers (Figure 2), 69 out of 810 participants (8.52%) were classified as indoor fallers. Two stratifying variables (TUG and FR) were identified. TUG was used as the first stratifying variable and the threshold was determined to be 18. Participants with TUG greater than or equal to 18 were found to have a higher fall rate of 18.25% compared with those with TUG less than 18, who had a lower fall rate of 6.73%. Among the participants with TUG greater than or equal to 18, the second stratifying variable (FR) was chosen and its threshold was set at 15. Participants with FR greater than or equal to 15 had a lower fall rate of 13.27% compared with those with FR less than 15, who had a fall rate of 61.54%.

Regarding the decision tree for outdoor fallers (Figure 3), 108 out of 807 participants (13.38%) were classified as out outdoor fallers, and two stratifying variables, namely FR and number of drugs taken, were identified. FR was chosen as the first stratifying variable and the

threshold was determined to be 17. Participants with FR less than 17 were found to have a higher fall rate of 24.64% compared with those with FR greater than or equal to 17, who had a relatively lower fall rate of 12.33%. Among those participants with FR less than 17, number of drugs taken was chosen as the second stratifying variable and its threshold was set to 2 for further discrimination. Those who took fewer than 2 drugs had a relatively low fall rate of 19.35% compared with those who took 2 or more drugs, who had a higher fall rate of 71.43%.

Random forest analysis with sampling method to deal with imbalanced classification

Table 2 shows the result of Random Forest for fallers with different sampling methods to treat imbalanced classification of fallers and non-fallers. Among the six sampling scenarios, the undersampling method showed the lowest accuracy rate of 55.75% (95% CI: 50.29%, 61.12%) whereas the other sampling methods showed an accuracy rate of more than 70%. Conversely, the sensitivity was
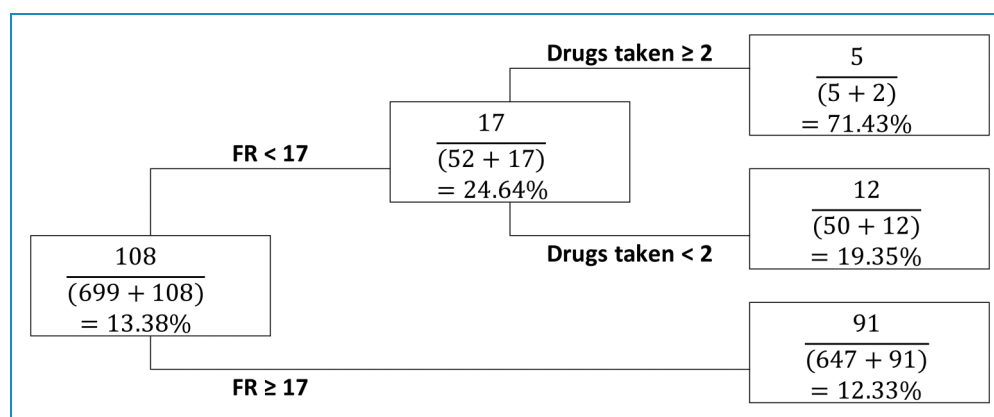
Figure 3. Decision tree for outdoor fallers
*overall accuracy rate = 85.76%.

Table 2. Results of random forest for fallers with different sampling methods to treat imbalanced classication of fallers and non-fallers.

	Oversampling method	Undersampling method	Both oversampling and undersampling method	ROSE function ($N = 1000$)	ROSE function ($N = 2000$)	ROSE function ($N = 5000$)
Accuracy (95% confidence interval)	77.29% (75.45%, 81.64%)	55.75% (50.29%, 61.12%)	71.09% (65.95%, 75.86%)	70.5% (65.33%, 75.31%)	72.27% (67.18%, 76.97%)	72.27% (67.18%, 76.97%)
Sensitivity	9.72%	56.94%	26.39%	38.89%	30.56%	18.06%
Specificity	95.5%	55.43%	83.15%	79.03%	83.52%	86.89%
Positive predictive value	36.84%	25.62%	29.69%	33.33%	33.33%	27.08%
Negative predictive value	79.69%	82.68%	80.73%	82.75%	81.68%	79.73%

Confusion matrix

Pre	True NF	True F
NF	255	65
F	12	7

Pre	True NF	True F
NF	148	31
F	119	41

Pre	True NF	True F
NF	222	53
F	45	19

Pre	True NF	True F
NF	211	44
F	56	28

Pre	True NF	True F
NF	223	50
F	44	22

Pre	True NF	True F
NF	232	59
F	35	13

True: true group membership; Pre: predicted group membership; NF: non-fallers; F: fallers.

Table 3. Results of random forest for indoor fallers with different sampling methods to treat imbalanced classication of fallers and non-fallers.

	Oversampling method	Undersampling method	Both oversampling and undersampling method	ROSE function ($N = 1000$)	ROSE function ($N = 2000$)	ROSE function ($N = 5000$)
Accuracy (95% confidence interval)	89.97% (86.27%, 92.95%)	60.47% (55.05%, 65.71%)	85.55% (81.34%, 89.11%)	76.70% (71.83%, 81.09%	76.40% (71.51%, 80.82%)	81.12% (76.54%, 85.15%)
Sensitivity	2.94%	55.88%	14.71%	32.35%	20.59%	8.82%
Specificity	99.67%	60.98%	93.44%	81.64%	82.62%	89.18%
Positive predictive value	50%	13.77%	20.0%	16.48%	11.67%	8.33%
Negative predictive value	90.21%%	92.54%	90.76%	91.54%	90.32%	89.77%

Confusion matrix:

Oversampling method

Pre	True NF	True F
NF	304	33
F	1	1

Undersampling method

Pre	True NF	True F
NF	186	15
F	119	19

Both oversampling and undersampling method

Pre	True NF	True F
NF	285	29
F	20	5

ROSE function ($N = 1000$)

Pre	True NF	True F
NF	249	23
F	56	11

ROSE function ($N = 2000$)

Pre	True NF	True F
NF	252	27
F	53	7

ROSE function ($N = 5000$)

Pre	True NF	True F
NF	272	31
F	33	3

True: true group membership; Pre: predicted group membership; NF: non-fallers; F: fallers.

the largest under the sampling method with 56.94% among all sampling scenarios, with the oversampling method carrying the lowest accuracy rate of 9.72%. On the other hand, the specificity value was the lowest for the undersampling method at 55.43%. Regarding the positive predictive value, the undersampling method showed the lowest value of 25.62% but a relatively higher negative predictive value of 82.68%.

Table 3 shows the result of Random Forest for indoor fallers with different sampling methods to treat imbalanced classification of fallers and non-fallers. The result was quite similar to that of fallers under Table 2. Again, the undersampling method showed the lowest accuracy rate of 60.47% (55.05%, 65.71%), the highest sensitivity of 55.88% and the lowest specificity of 60.98% compared with other sampling methods. Regarding the positive predictive values, the undersampling method showed a relatively smaller value of 13.77% compared with oversampling method (50%) whereas it showed a maximum value of 92.54% compared with other sampling scenarios.

Table 4 shows the result of Random Forest for outdoor fallers with different sampling methods to treat imbalanced classification of fallers and non-fallers. Again, the result was consistent with that of fallers (Table 2) and indoor fallers (Table 3). The undersampling method showed the lowest accuracy rate of 54.57% (49.10%, 59.96%), the

highest sensitivity of 52.94%, the lowest specificity of 54.86% compared with other sampling methods. Regarding the positive predictive values, the undersampling method showed the second lowest value of 17.20% whereas it showed the maximum value of 86.61% compared with other sampling scenarios.

Discussion

We identified the predictive attributes for fallers generally, and indoor and outdoor fallers specifically, using decision tree analysis for accidental falls in community-dwelling older people in Hong Kong. A clinical algorithm derived from decision tree analysis using the TUG, the FR test, BMI, high BP, osteoporosis, and a number of drugs taken was developed for fall screening for community-dwelling older people which could be used for developing a clinical algorithm for fall screening.

The summary of the cross-validation process with the same split of training and testing dataset for 10 training iterations. found that TUG, FR, BMI, high BP, and number of drugs taken could be used as factors in risk assessment and fall screening for community-dwelling older people. Age, gender, level of education, previous injury, comorbidity, the performance of IADL, mobility status, use of walking aid, social support, and visual

Table 4. Results of random forest for outdoor fallers with different sampling methods to treat imbalanced classication of fallers and non-fallers

	Oversampling method	Undersampling method	Both oversampling and undersampling method	ROSE function ($N = 1000$)	ROSE function ($N = 2000$)	ROSE function ($N = 5000$)
Accuracy (95% confidence interval)	84.07% (79.73%, 87.8%)	54.57% (49.10%, 59.96%)	76.11% (71.20%, 80.55%)	77.29% (72.45%, 81.64%)	76.99% (72.14%, 81.37%)	79.35% (74.65%, 83.53%)
Sensitivity	5.88%	52.94%	9.80%	23.53%	21.57%	13.73%
Specificity	97.92%	54.86%	87.85%	86.81%	86.81%	90.97%
Positive predictive value	33.33%	17.20%	12.50%	24.0%	22.45%	21.21%
Negative predictive value	85.46%	86.81%	84.62%	86.51%	86.21%	85.62%
Confusion matrix						

Confusion matrices:

Oversampling method:

Pre	True NF	True F
NF	282	48
F	6	3

Undersampling method:

Pre	True NF	True F
NF	158	24
F	130	27

Both oversampling and undersampling method:

Pre	True NF	True F
NF	253	46
F	35	5

ROSE function ($N = 1000$):

Pre	True NF	True F
NF	250	39
F	38	12

ROSE function ($N = 2000$):

Pre	True NF	True F
NF	250	40
F	38	11

ROSE function ($N = 5000$):

Pre	True NF	True F
NF	262	44
F	26	7

All the confusion matrices have been shown for reference under Figures 4 to 6.

acuity were not identified as stratifying variables in the decision tree models.

Our findings can only be applied to those older people who are able to ambulate independently in the community. Although self-reporting of fall events has been criticized due to the tendency to under-report actual fall events, the clinical value of this retrospective self-reported method in reflecting the actual conditions of falls is demonstrated by the fact that it is still being widely used in fall studies.[26] A systematic review reports that recall of any fall in the previous year is relatively specific (91%–95%) but somewhat less sensitive (80%–89%) than intensive prospective data collection. However, 12-month recall is still better than recall for falls in the previous 3 or 6 months.[14] Screening for falls and fall risk is aimed at preventing or reducing fall risk; persons screened into a high-risk group warrant further evaluation.[15] We believe that self-reporting of recall of fall events was useful in fall detection for preventing or reducing fall risk. It is recommended that all older people who are under the care of a health professional (or a caregiver) should be asked at least once a year about falls.[27]

The clinical algorithm we found for TUG of fallers (cutoff range $\geq$ 15–18; including both indoor and outdoor fallers) was consistent with the reference framework for preventive care for older people in primary care settings in the local community.[28] According to the updated guideline, the cut-off value of 15 s is conventionally adopted to discriminate fallers and non-fallers in healthy, highly functional older people.[7] However, other studies have found that TUG is also a sensitive and specific measure of frailty[29] but might not be able to reliably identify pre-frail individuals.[30]

Measurement of unsteadiness in balance, as represented by FR, was also consistent with the literature, with a reported cutoff range of 15–17 for various fallers in this study. FR is the most common tool in risk assessment for older people living independently in their own homes and in community dwellings, and has reached sensitivity and specificity values greater than 70. It has been tested in other studies with different populations, including recurrent fallers and cognitively impaired persons.[8]

Conventionally, the risk of falls is thought to increase for those taking 4 or more medications.[31] Our results suggested a number of drugs taken $\geq$ 2 for a stringent screening of fall risk (slightly different from the risk factor of polypharmacy). The finding regarding high BP might also be associated with postural hypotension, which is a risk factor for falls in community-living older people.[32] It was interesting to find that BMI > 22 led to greater risk of falls, as BMI = 22 is in the normal range for older persons, but it might imply that obesity was associated with a higher prevalence of falls and stumbling during ambulation.[32] Our finding of BMI >22 with a higher risk of fall was consistent with

another decision tree study based on a prospective observational multicenter study.[33] Another study has also found that in community-dwelling older people, obese individuals had a 31% higher risk of having fallen, and among them, a 41% higher risk of using 4 or more medications which might be associated with a higher risk of certain physical comorbidites commonly associated with weight gain, including diabetes, heart disease and high BP.[34]

There have been pilot studies for older people using machine-learning techniques to generate a decision tree model that could be used to predict falls for the medical staff in the clinical settings and medical fields.[33,35] Our findings might be useful in developing screening tools for fall risk of community-living older people and may have an impact beyond Hong Kong – the use of decision tree analysis for clinical algorithms for accidental falls in community-dwelling older people minimizes the time required for processing a big dataset and creates decision-making patterns, which also paves the way for utility-based decision making using supervised ML in fall risk detection for the older people in primary care settings in future. By the same token, it is not difficult to identify current research emphasize the use of ML in predicting fall incidence.[11,12,36] Another benefit of a simple decision tree is that the model itself is easy to interpret and visualize. It also requires low computation, thus reducing the time to implement but carrying low accuracy compared with other advanced MI methods such as random forest etc. However, the clinician can know which variable and which value of the variable uses to split the data, predicting the faller outcome status quickly. Compared with a single decision tree, random forest is an advanced classification algorithm consisting of many decision trees combined to get a more accurate and precise result. It can also avoid and prevent overfitting by using multiple trees. Nonetheless, random forest consumes more computation time in the process of generating and analyzing. In case of all these reasons, decision tree analysis was adopted for preliminary reference in this manuscript instead of random forest algorithms. To pursue a predictive model with a higher accuracy rate and sensitivity that can be used in practice, more advanced ML methods such as random forest have been used in our study combined with different sampling methods such as oversampling, undersampling, both oversampling and undersampling as well as ROSE function. No matter which type of fallers (fallers, indoor fallers and outdoor fallers) were predicted under the random forest algorithm, it was found that the undersampling method showed the highest sensitivity (56.94%, 55.88% and 52.94%) among all the sampling methods although the accuracy rates were relatively lower (55.75%, 60.47% and 54.57%) when compared with other sampling methods. Due to this reason, the undersampling method could be considered to be the most reliable sampling method to deal with imbalance classification of falling incidence. More further research studies on different sampling scenarios on ML algorithms warrant investigation. Moreover, some advanced ML method such as support vector machine, etc. could be considered to be incorporated in future studies.

Our study had limitations. All participants were recruited in a primary care setting who only showed up for once, it would be difficult for us to conduct a lengthy comprehensive assessment during screening and a follow-up assessment for them after 1-year. Second, we did not take into account recurrent fallers because the incidence was small in the sample and the prediction model might not be accurate. However, the aim of our current study is to provide a clinical guide for supervised learning in ML, and the predictors have already been identified as potential predictive factors in the literature. Clinically, our decision trees are consistent with the evidence and the accuracy rates based on all the decision trees contribute to the model-fit. However, there might be variation due to sample bias and this may be due to inadequate power level.

Third, the sample size was comparatively small in terms of ML. In fact, it is hard to determine the minimum sample size of performing ML method for this type of data as it depends on a vast number of factors such as complexity of the problem, complexity of the learning algorithm etc. For example, large neural networks can be trained on billions of samples[37] while simple ML models can use only a few samples.[38] Fourth, the participants were healthy older adults resided in community dwellings, the algorithms developed in this study might not be useful to predict falls for the high-risk groups or those living in residential institutions. Moreover, cultural differences and national healthcare support structures can influence conclusions in this study that are not necessarily applicable in other countries.

Conclusion

We identified the predictive attributes for fallers generally, and indoor and outdoor fallers based on a fast screening assessment, specifically, using decision tree analysis for accidental falls in community-dwelling older people in a primary care setting, which could be used for developing a clinical algorithm for fall screening in the community. Our findings might be useful to assist clinicians in decision-making, which also paves the way for the future use of ML algorithms in fall risk detection in primary care settings.

ORCID iD: Kenneth N K Fong https://orcid.org/0000-0001-5909-4847

References

1. HealthyHK – Public Health Information and Statistics of Hong Kong. *Life expectancy*. The Government of the Hong Kong SAR. Retrieved 10 January 2023 from https://www.healthyhk.gov.hk/phisweb/en/chart_detail/13/

2. Dodwell D. *Unlike in New York, Hong Kong's high-rise living has helped its communities thrive*. South China Morning Post April 2022. Retrieved 10 January 2023 from https://www.scmp.com/comment/opinion/article/3177458/unlike-new-york-hong-kongs-high-rise-living-has-helped-its

3. The Census and Statistics Department. *Population and household statistics analysed by District Council District* 2021. The Government of the Hong Kong SAR Retrieved 10 January 2023 from http://www.censtat.gov.hk

4. Fong KNK, Siu AMH, Au Yeung K, et al. Falls among the community-living elderly people in Hong Kong: a retrospective study. *Hong Kong J Occup Ther* 2011; 21: 33–40.

5. Bergen G, Stevens MR and Burns ER. Falls and fall injuries among adults aged ≥65 years — United States, 2014. *MMWR Morb Mortal Wkly Rep* 2016; 65: 993–998.

6. Ganz DA and Latham NK. Prevention of falls in community dwelling older adults. *N Engl J Med* 2020; 382: 734–743.

7. Montero-Odasso M, van der Velde N, Martin F,C, et al. World guidelines for falls prevention and management for older adults: a global initiative. *Age Ageing* 2022; 51: 1–36.

8. Scott V, Votova K, Scanlan A, et al. Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings. *Age Ageing* 2007; 36: 130–139.

9. Fortinsky RH, Iannuzzi-Sucich M, Baker DI, et al. Fall-risk assessment and management in clinical practice: views from healthcare providers. *J Am Geriatr Soc* 2004; 52: 1522–1526.

10. Hill KD, Moore KJ, Dorevitch MI, et al. Effectiveness of falls clinics: an evaluation of outcomes and client adherence to recommended interventions. *J Am Geriatr Soc* 2008; 56: 600–608.

11. Lathouwers E., Dillen A., Díaz M.A, et al. Characterizing fall risk factors in Belgian older adults through machine learning: a data-driven approach. *BMC Public Health* 2022; 22: 2210..

12. Rafiq M, McGovern A, Jones S, et al. Falls in the elderly were predicted opportunistically using a decision tree and systematically using a database-driven screening tool. *J Clin Epidemiol* 2014; 67: 877–886.

13. Chiu HFK, Lee HC, Chung WS, et al. Reliability and validity of the cantonese version of mini-mental state examination: a preliminary study. *J Hong Kong College Psych* 1994; 4: 25–28.

14. Ganz DA, Higashi T and Rubenstein LZ. Monitoring falls in cohort studies of community-dwelling older people: effect of the recall interval. *J Am Geriatr Soc* 2005; 53: 2190–2194.

15. Census and Statistics Department. *Population and household statistics analysed by District Council District 2012*. Hong Kong: the Census and Statistics Department 2022. Accessed February 14, 2014 from http://www.censtatd.gov.hk/hkstat/sub/sp150.jsp?productCode=B1130301

16. World Health Organisation. *WHO global report on falls prevention in older age*. France: WHO Press; 2007.

17. Podsiadlo D and Richardson S. The timed "up & go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991; 39: 142–148.

18. Duncan PW, Weiner DK, Chandler J, et al. Functional reach: a new clinical measure of balance. *J Gerontol Med Sci* 1990; 45: M192–M197.

19. Hetherington R. The snellen chart as a test of visual acuity. *Psychol Forsch* 1954; 24: 349–357.

20. Rockwood K, Awalt E, Carver D, et al. Feasibility and measurement properties of the functional reach and the timed up and go tests in the Canadian study of health and aging. *J Gerontol A Biol Sci Med Sci* 2000; 55: M70–M73.

21. Shumway-Cook A, Brauer S and Woollacott M. Predicting the probability for falls in community-dwelling older adults using the timed up & go test. *Phy Ther* 2000; 80: 896–903

22. Giorgetti MM, Harris BA and Jette A. Reliability of clinical balance outcome measures in the elderly. *Phys Ther Res Int* 1998; 3: 274–283.

23. Terry M, Elizabeth J and Mayo F. *An introduction to recursive partitioning using the rpart routines*. Accessed October 21, 2022 from https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf

24. Nguyen QH, Ly H, Ho LS, et al. Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Math Probl Eng* 2021: 4832864.

25. Lunardon N, Menardi G and Torelli N. ROSE: a package for binary imbalanced learning. *R Journal* 2014; 6: 82–92.

26. Mackenzie L, Byles J.D and Este C. Validation of self-reported fall events in intervention studies. *Clin Rehabil* 2006; 20: 331–339.

27. Panel on prevention of falls in older persons, American Geriatrics Society and British Geriatrics Society. Summary of the updated American geriatrics society/British geriatrics society clinical practice guideline for prevention of falls in older persons. *J Am Geriatr Soc* 2011;59:148–157.

28. Food and Health Bureau, Government Secretariat. *Hong Kong reference framework for preventive care for older adults in primary care settings: module on falls in elderly* 2015. The Government of the Hong Kong Special Administrative Region; www.pco.gov.hk. (Accessed 15 Oct 2019).

29. Clegg A, Rogers L and Young J. Diagnostic test accuracy of simple instruments for identifying frailty in community-dwelling older people: a systematic review. *Age Ageing* 2015; 44: 148–152.

30. Sava GM, Donoghue OA, Horgan F, et al. Using timed up-and-go to identify frail members of the older population. *J Gerontol A Biol Sci Med Sci* 2013; 68: 441–446.

31. Hammond T and Wilson A. Polypharmacy and falls in the elderly: a literature review. *Nurs Midwifery Stud* 2013; 2: 171–175.

32. Fjeldstad C, Fjeldstad AS, Acree LS, et al. The influence of obesity on falls and quality of life. *Dyn Med* 2008; 7: 4..

33. Deschamps T,Le, Goff CG, Berrut G, et al. A decision model to predict the risk of the first fall onset. *Exp Gerontol* 2016; 81: 51–55.

34. Mitchell RJ, Lord SR, Harvey LA, et al. Associations between obesity and overweight and fall risk, health status and quality of life in older people. *Aust NZ J Public Health* 2014; 38: 13–18.

35. Gillain S, Boutaayamou M, Schwartz C, et al. 2019. Using supervised learning machine algorithm to identify future fallers based on gait patterns: a two-year longitudinal study. *Exp Gerontol* 2019; 127. doi.org/10.1016/j.exger.2019.110730

36. Ojetola OE, Gaura I and Brusey J. Fall detection with wearable sensors–safe (smart fall detection). 2011 Seventh international conference on intelligent environments, Nottingham, UK, 2011, 318–321.

37. Hoffmann J, Borgeaud S, Mensch A, et al. *Training Compute-Optimal Large Language Models*. arXiv, Mar. 29, 2022. Cornell University: Computer Science. Accessed 13 Feb 2023 from

38. Zhang Y and Ling C. A strategy to apply machine learning to small datasets in materials science. *NPJ Comput Mater* 2018; 4: 25.