Classification of Fall Types in Parkinson's Disease From Self-report Data Using Natural Language Processing

Jeanne M. Powell[1][0000-0002-3494-2376], Yuting Guo[1][0000-0002-8919-0888], Abeed Sarker[1,2][0000-0001-7358-544X], and J. Lucas McKay[1,2][0000-0002-8361-8943]

[1] Emory University, Atlanta GA 30322, USA
[2] Georgia Institute of Technology, Atlanta GA 30332, USA
{jeanne.marie.powell,yuting.guo,abeed.sarker,j.lucas.mckay}@emory.edu

Abstract. Falls are a leading cause of injury globally, and people with Parkinson's disease are particularly at risk. An important step in reducing the probability of falls is to identify their causes, but manually classifying fall types is laborious and requires expertise. Natural language processing (NLP) approaches hold potential to automate fall type identification from descriptions. The aim of this study was to develop and evaluate NLP–based methods to classify fall types from Parkinson's disease patient self-report data. We trained supervised NLP classifiers using an existing dataset consisting of both structured and unstructured data, including the age, gender, and duration of Parkinson's disease of the faller, as well as the fall location, free-text fall description, and fall class of each fall. We trained supervised classification models to predict fall class based on these attributes, and then performed an ablation study to determine the most important factors influencing the model. The best performing classifier was a hard voting ensemble model that combined the Adaboost, unweighted decision tree, weighted k-nearest neighbor, naïve Bayes, random forest, and support vector machine classifiers. On the testing set, this ensemble classifier achieved an F_1-macro of 0.89. We also experimented with a transformer-based model, but its performance was subpar compared to that of the other models. Our study demonstrated that automatic fall type classification in Parkinson's disease patients is possible via NLP and supervised classification.

Keywords: Falls, Natural Language Processing, Text Classification, Parkinson's disease.

1 Introduction

Falls are unintentional events where a person lands on a lower level [1, 2], which can result in significant personal, financial, and health costs [3–6]. For example, falls were the leading cause of injury-related death in the United States between 2007 and 2016 [3]. These high costs could be minimized with a better understanding of the causes of falls and subsequent implementation of preventative measures.

Many studies report overall fall frequency without accounting for the circumstances surrounding a fall, which limits our understanding of their etiology [7]. Falls are heterogeneous and can result from multiple types of biomechanical perturbations, including perturbations to an individual's base of support (BoS; e.g., trips) or center of mass (CoM; e.g., overextension during bending) [2]. BoS falls are more common in healthy older adults compared to CoM falls [2]. However, the opposite holds true in subpopulations of people, such as people with Parkinson's disease, where disease-related postural instability results in more CoM falls [1].

People with Parkinson's disease are more likely to fall and be frequent fallers than healthy older adults [1, 6]. Falling in this population can be incapacitating, often resulting in soft tissue injuries, and disabling even early in disease progression [1]. Therefore, it is of particular importance to predict and prevent falls in this population.

A necessary step in this pursuit is to track falls and fall circumstances because risk factors for trips and slips might differ from those for falls due to impaired self-motion or other causes. To determine the cause of a fall, one must collect free-text information about the circumstances of the fall from the faller [6]. Historically, fall classes have been manually coded from these free-text descriptions [6, 8–10], but this practice is subjective, resource intensive, and difficult to scale. Recent advances in the field of natural language processing (NLP) hold exciting promise to automate processes such as fall type classification from free-text fall descriptions.

NLP techniques have been used in many biomedical domains, including mining unstructured electronic health records [11]. For example, Tohira and colleagues trained support vector machine (SVM) and random forest (RF) classifiers to detect falls from ambulance services provider reports [12]. Electronic health records and patient self-reports provide rich data that can capture nuances that structured medical data may miss [13]. NLP techniques may aid in Parkinson's disease diagnosis, given its impact on language production. Pérez-Toro and colleagues demonstrated that NLP techniques could be leveraged to distinguish people with Parkinson's disease from healthy older adults based on differences transcribed in speech patterns [14]. Given that Parkinson's disease is the second most common neurodegenerative disease worldwide, affecting over six million people globally [15], it is extremely important to gain a nuanced understanding of the disease.

Here, we aimed to develop an NLP classification model to distinguish CoM falls from other fall types in people with Parkinson's disease based on patient-provided free-text descriptions. Our particular focus is on CoM falls as they are more common in people with Parkinson's disease [1].

2 Methods

2.1 Fall Self-report Dataset

In a recent study [16], we followed patients with Parkinson's disease for 12-months to correlate fall risk with biomarkers of balance control. Participants tracked falls on monthly "fall calendars" and missing data and fall details were acquired over telephone interviews. Our dataset consisted of 124 fall self-reports collected from 23 individuals. Demographic information about those individuals is provided in Table 1.

Each fall self-report included structured data (i.e., age, gender, and time since Parkinson's disease diagnosis) and unstructured data (free-text description of the fall and its location). Dr. McKay, an expert in the biomechanics of falls, classified fall causes (i.e, CoM, BoS, or 'Other') based on fall descriptions, and Powell classified falls as occurring inside the home or not using location descriptions. Three fall descriptions were modified because the patient described the fall as "same as fall #1". In these cases, the full description used for #1 was copied. The fall descriptions were also manually checked for spelling errors.

The average length of the fall descriptions was 38 words, with the shortest consisting of only 3 words and the longest of 170 words. There were 922 unique words in the dataset before processing and 731 unique words after processing (i.e., stemming and removal of stop words).

2.2 Data Pre-processing

All binary categorical variables (i.e., gender, fall class, and fall location) were one-hot encoded. Our numerical factors, age and disease duration, were scaled. Free-text fall descriptions were pre-processed by lowercasing all text, removing English stop words and punctuations, and tokenizing the remaining text. Each word token was stemmed using the Porter stemmer [17]. Fall descriptions were then vectorized as follows. We generated two sets of features from the pre-processed description texts—n-grams and word clusters. A word n-gram is a sequence of contiguous n words in a text segment. This feature enabled us to represent a document using the union of its terms. We used 1-, 2-, and 3-grams as features with the max number of features set to 150. The n-grams were vectorized so that each n-gram was represented by a numeric value in the feature vector indicating its frequency within a given instance. To enable a more generalized representation of the terms, we used the CMU word clusters [18]. The word clusters were generated via a two-step process—dense vector representations of words were first learned from large unlabeled data using the method described in [18] so that similar terms were close together in vector space, and then the words were grouped via hierarchical clustering. The word clusters were represented as unigrams during training.

2.3 Fall Class Classification

We modeled the discrimination between CoM- and Other-class falls as a binary classification problem, using both structured and unstructured features. Because the dataset was relatively small, we applied a predefined 3-fold cross validation for (80% of data) training and (20% of data) evaluation. We experimented with multiple classifiers, specifically: naïve Bayes (NB), K-Nearest Neighbors (KNN), SVM, RF, Adaboost with single split trees as base classifiers, a Decision Tree (DT) classifier, and a hard-voting ensemble classifier with contributions from each of the previously mentioned classifiers. We experimented with both weighted and unweighted KNN and DT classifiers to account for our unbalanced classes. The performances of the classifiers were compared using the F1-macro score on the test data because that metric is more appropriate when classes are unbalanced. We then performed an ablation study to determine the individual impact of each factor on model performance, as well as model performance when only trained on the text.

We also modeled the discrimination between CoM- and Other-class falls using the RoBERTa transformer model [19]. For this experiment, we utilized the unprocessed, free-text fall descriptions to predict fall-class labels and did not include other factors in the model. We applied a 3-fold cross validation for (80% of data) training and (20% of data) evaluation. The model was trained for 2, 5, and 10 epochs. All model parameters were fine-tuned during training. Performance was measured by taking the median of the F_1-macro score.

Confidence intervals were calculated using bootstrapping with samples taken from the test prediction and ground truth datasets with replacement over 1000 iterations.

3 Results

3.1 Demographic Information

Our fall dataset included 23 individuals with an average age of 67.3 years and an average Parkinson's disease duration of 8.9 years (Table 1). There was no significant difference between the overall fall frequency by gender (p = 0.203), nor in fall frequency between gender within each fall class (CoM: p = 0.267; BoS: p = 0.662; Other: p = 0.614).

Table 1. Demographic information and fall frequency of patients

	Overall	Women	Men	P-Value
n	23	8	15	
Age, mean (SD)	67.3 (7.1)	65.4 (6.4)	68.3 (7.5)	0.333
Duration, mean (SD)	8.9 (5.1)	8.6 (5.5)	9.1 (5.1)	0.838
Total, mean (SD)	5.4 (4.9)	7.5 (6.1)	4.3 (4.0)	0.203
CoM, mean (SD)	3.8 (4.5)	5.5 (5.5)	2.9 (3.8)	0.267
BoS, mean (SD)	1.1 (1.0)	1.2 (1.5)	1.0 (0.7)	0.662
Other, mean (SD)	0.5 (1.5)	0.8 (2.1)	0.3 (1.0)	0.614

Of the 124 unique falls, 88 falls occurred due to perturbations to the individuals' center of mass (CoM), 25 falls occurred due to perturbations to the individuals' base of support (BoS), and the remaining 11 falls occurred for other reasons, including falling during exercise (n = 9), low blood pressure (n =1), and rolling out of bed (n = 1). Because of the relatively low number of BoS- and Other-class falls in our dataset, we collapsed the data into CoM- and Other-class falls. There were 88 CoM-class falls and 36 Other-class falls.

3.2 Binary Classification Hyperparameter Selection

For each classifier, we iterated through a wide range of hyper-parameter values and calculated the mean squared error between the actual fall class and predicted fall class at each hyperparameter value. Mean squared error is inversely proportional to model performance. Therefore, the hyperparameter value that resulted in the lowest mean squared error for each model was chosen (Table 2).

3.3 Binary Classifier Performance

Each classifier was trained on the same 80% of the dataset using 3-fold cross validation with the same folds. Then, each classifier was tested on the same 20% of data that had been excluded from the training phase, and we report the resulting F_1-macro score and its 95% confidence interval (CI). The ensemble classifier had the highest performance out of all trained classifiers (F_1-macro = 0.89, 95% CI: [0.67-1]; Table 2).

Adaboost. We determined the optimal number of estimators for the Adaboost model to be 27 by training on all values of n between 1 and 100, inclusive, and selecting the value of n with the lowest mean squared error between the true and predicted fall classes. The model achieved an F_1-macro of 0.80 (95% CI: [0.55-0.96]; Table 2).

Decision Tree. We trained two DT classifiers, one with weights set to the inverse frequency of each class in the training set (DTa) and the other with equal weights for both classes (DTb). We determined the optimal maximum depth of DTa to be 23 and the optimal maximum depth of DTb to be 18. The weighted model DTa achieved an F_1-macro of 0.67 (95% CI: [0.45-0.85]; Table 2). The unweighted model DTb performed better, achieving an F_1-macro of 0.72 (95% CI: [0.50-0.90]; Table 2).

K-Nearest Neighbor. We trained two KNN classifiers, one with the weight function equal to 'distance' (KNNa) and the other with the weight function equal to 'uniform' (KNNb). For the weighted model, KNNa, the optimal value of k was 6 and for the unweighted model, KNNb, the optimal value of k was 10. The weighted model KNNa achieved an F_1-macro of 0.84 (95% CI: [0.66-1; Table 2). The unweighted model KNNb achieved an F_1-macro of 0.78 (95% CI: [0.53-0.95]; Table 2).

Naïve Bayes. We trained a Gaussian Naïve bayes classifier. The model achieved an F_1-macro of 0.84 (95% CI: [0.63-1]; Table 2).

Random Forest. We determined the optimal number of estimators for the RF model to be 25 by training each model on all values of n between 1 and 100, inclusive, and selecting the value of n that resulted in the lowest mean squared error between the true fall classes and predicted fall classes. The RF model achieved an F_1-macro of 0.78 (95% CI: [0.53-0.95]; Table 2).

Support Vector Machine. We trained the SVM model with gamma set to ‘scale’ and kernel set to ‘rbf’ on all values of C between 1 and 100 inclusive. We found that the mean squared error between the true and predicted fall class was at its lowest when C was 4. The model achieved an F_1-macro of 0.78 (95% CI: [0.53-0.95]; Table 2).

Ensemble. The ensemble was composed of one of each type of the classifiers described above. The weighted KNN model KNNa and unweighted DT model DTa were included in the ensemble because they outperformed their complementary model. We set the voting for the ensemble equal to hard. The ensemble achieved an F_1-macro of 0.89 (95% CI: [0.67-1]; Table 2).

Table 2. Classifier performance at predicting fall type

Classifier	Hyperparameters	F_1-macro	95% CI
Adaboost	n_estimators = 27	0.80	0.55-0.96
DTa	max_depth = 23, class_weight = {0:70.0, 1:29.0)	0.67	0.45-0.85
DTb	max_depth = 18	0.72	0.50-0.90
KNNa	k = 6, weights = ‘distance’	0.84	0.66-1
KNNb	k = 10, weights = ‘uniform’	0.78	0.53-0.95
NB	N/A	0.84	0.63-1
RF	n_estimators = 25	0.78	0.53-0.95
SVM	Gamma = ‘scale’, kernel = ‘rbf’, C = 4	0.78	0.53-0.95
Ensemble	{NB, KNNa, SVM, RF, Adaboost, DTb} voting = ‘hard’	**0.89**	**0.67-1**

3.4 Ablation Study of Features in the Ensemble Model

Table 3. Results of ablation study on the ensemble model

Dropped Feature	F_1-macro	95% CI
Sex	0.80	0.59-0.96
Location	0.84	0.66-1
PD Duration	0.80	0.59-0.95
Age	**0.89**	**0.70-1**
Word Clusters	0.82	0.63-0.96
1,2,3-grams	0.78	0.55-0.95
Sex, Location, PD Duration, & Age	0.83	0.58-1

We chose to perform an ablation study on the ensemble model because it achieved the highest F_1-macro score out of all trained classifiers. The F_1-macro score of the model decreased with the removal of each factor, except age (Table 3).

3.5 Binary Classification Using RoBERTa

The RoBERTa transformer model was trained on 80% of the dataset using 3-fold cross validation. Classifier performance was measured using the F_1-macro metric. The RoBERTa model had equal performance across epochs (F_1-macro = 0.42; Table 4).

Table 4. Performance of RoBERTa model

Epochs	F_1-macro	95% CI
2	0.42	0.37-0.46
5	0.42	0.37-0.45
10	0.42	0.37-0.45

We also trained the machine learning models on the vector representations generated by RoBERTa using the same hyperparameters in Table 2. The results show that using RoBERTa as a feature generator underperformed compared to using the n-grams and word clusters as features (Table 5).

Table 5. Classifier performance using RoBERTa-generated vector representations

Model	F_1-macro	95% CI
Adaboost	0.61	0.40-0.81
DTa	0.70	0.49-0.88
DTb	0.58	0.38-0.77
KNNa	0.71	0.44-0.90
KNNb	0.44	0.37-0.48
NB	0.48	0.29-0.66
RF	0.61	0.39-0.81
SVM	0.41	0.34-0.46

4 Discussion

We trained multiple machine learning classifiers to perform a binary classification task to categorize falls as CoM- or Other-type falls based on patient-provided free-text descriptions of the circumstances surrounding each fall. We found that a hard-voting ensemble classifier performed better than individual classifiers and transformer-based models on this task, achieving an F_1-macro of 0.89 (95% CI: [0.67-1]). This serves as a proof of concept that the historically resource intensive task of manual fall type classification can be automated using NLP models. Importantly, our ensemble approach obtained high performance despite the relatively small size of the annotated dataset, which is often the limiting factor for supervised classification tasks.

Current clinical best practices for tracking falls in Parkinson’s disease involve retrospective patient reports during regular Neurologist visits, which may lead to recall bias. Although standardized instruments exist [20], they are very uncommon in clinical practice due to the burden on patients and providers. A technology for classification of falls circumstances and causes based on patient reports may enable future “online trials” and reduce misclassification errors in research studies, which might contribute to the high variability across studies applying exercise on fall risk in Parkinson's disease [21].

4.1 Better Performance From ML Classifiers

Unexpectedly, traditional machine learning classifiers and their ensemble outperformed the RoBERTa transformer model in our study. We initially expected the transformer-based architecture to perform better, but it categorized all falls as CoM falls, suggesting that the small size and class imbalance of our data were not ideal for this model. Furthermore, our ablation study revealed that non-text factors were important contributors to model performance, suggesting that using only free-text descriptions may have hindered the transformer model’s performance.

4.2 A Lack of Gender Differences in Fall Frequency and Type

In people with Parkinson’s disease, there is no gender difference in fall frequency, unlike healthy older adults [1]. Our study found no difference in falls between genders. Future research could explore if the circumstances surrounding falls differ by gender in people with Parkinson’s disease, similar to healthy older adults [22].

4.3 Future Work

To improve our study’s generalizability, we plan to expand our small, unbalanced dataset with limited vocabulary by scraping social media profiles of people with Parkinson’s disease. This will allow us to test and retrain our ensemble model for better performance. We also plan to explore other transformer-based models and strategies for integrating structured data.

5 Conclusion

Our study demonstrated that it is possible to automate the laborious process of fall type identification by using supervised classification methods that integrate structured and unstructured data. An ensemble classification approach produced excellent results, outperforming a state-of-the-art transformer model despite the small size of annotated data. Find the dataset and code at the following repository: https://github.com/jeannempowell/PD_Falls_NLP.

References

1. Bloem, B.R., Grimbergen, Y.A.M., Cramer, M., Willemsen, M., Zwinderman, A.H.: Prospective assessment of falls in Parkinson's disease. J. Neurol. 248, 950–958 (2001). https://doi.org/10.1007/s004150170047.
2. Maki, B.E., Holliday, P.J., Topper, A.K.: A Prospective Study of Postural Balance and Risk of Falling in An Ambulatory and Independent Elderly Population. J. Gerontol. 49, M72–M84 (1994). https://doi.org/10.1093/geronj/49.2.M72.
3. Burns, E., Kakara, R.: Deaths from Falls Among Persons Aged ≥65 Years — United States, 2007–2016. MMWR Morb. Mortal. Wkly. Rep. 67, 509–514 (2018). https://doi.org/10.15585/mmwr.mm6718a1.
4. Florence, C.S., Bergen, G., Atherly, A., Burns, E., Stevens, J., Drake, C.: Medical Costs of Fatal and Nonfatal Falls in Older Adults: Medical Costs of Falls. J. Am. Geriatr. Soc. 66, 693–698 (2018). https://doi.org/10.1111/jgs.15304.
5. Haddad, Y.K., Bergen, G., Florence, C.S.: Estimating the Economic Burden Related to Older Adult Falls by State. J. Public Health Manag. Pract. 25, E17–E24 (2019). https://doi.org/10.1097/PHH.0000000000000816.
6. Stack, E., Ashburn, A.: Fall events described by people with Parkinson's disease: implications for clinical interviewing and the research agenda. Physiother. Res. Int. 4, 190–200 (1999). https://doi.org/10.1002/pri.165.
7. Ross, A., Yarnall, A.J., Rochester, L., Lord, S.: A novel approach to falls classification in Parkinson's disease: development of the Fall-Related Activity Classification (FRAC). Physiotherapy. 103, 459–464 (2017). https://doi.org/10.1016/j.physio.2016.08.002.
8. Ashburn, A., Stack, E., Ballinger, C., Fazakarley, L., Fitton, C.: The circumstances of falls among people with Parkinson's disease and the use of Falls Diaries to facilitate reporting. Disabil. Rehabil. 30, 1205–1212 (2008). https://doi.org/10.1080/09638280701828930.
9. Pelicioni, P.H.S., Menant, J.C., Latt, M.D., Lord, S.R.: Falls in Parkinson's Disease Subtypes: Risk Factors, Locations and Circumstances. Int. J. Environ. Res. Public. Health. 16, 2216 (2019). https://doi.org/10.3390/ijerph16122216.
10. Magnani, P.E., Genovez, M.B., Porto, J.M., Zanellato, N.F.G., Alvarenga, I.C., Freire, R.C., de Abreu, D.C.C.: Use of the BESTest and the Mini-BESTest for Fall Risk Prediction in Community-Dwelling Older Adults Between 60 and 102 Years of Age. J. Geriatr. Phys. Ther. 43, 179–184 (2020). https://doi.org/10.1519/JPT.0000000000000236.
11. Houssein, E.H., Mohamed, R.E., Ali, A.A.: Machine Learning Techniques for Biomedical Natural Language Processing: A Comprehensive Review. IEEE Access. 9, 140628–140653 (2021). https://doi.org/10.1109/ACCESS.2021.3119621.
12. Tohira, H., Finn, J., Ball, S., Brink, D., Buzzacott, P.: Machine learning and natural language processing to identify falls in electronic patient care records from ambulance attendances. Inform. Health Soc. Care. 47, 403–413 (2022). https://doi.org/10.1080/17538157.2021.2019038.

13. Guetterman, T.C., Chang, T., DeJonckheere, M., Basu, T., Scruggs, E., Vydiswaran, V.V.: Augmenting Qualitative Text Analysis with Natural Language Processing: Methodological Study. J. Med. Internet Res. 20, e231 (2018). https://doi.org/10.2196/jmir.9702.
14. Pérez-Toro, P.A., Vásquez-Correa, J.C., Strauss, M., Orozco-Arroyave, J.R., Nöth, E.: Natural Language Analysis to Detect Parkinson's Disease. In: Ekštein, K. (ed.) Text, Speech, and Dialogue. pp. 82–90. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-27947-9_7.
15. Dorsey, E.R., Elbaz, A., Nichols, E., Abbasi, N., Abd-Allah, F., Abdelalim, A., Adsuar, J.C., Ansha, M.G., Brayne, C., Choi, J.-Y.J., Collado-Mateo, D., Dahodwala, N., Do, H.P., Edessa, D., Endres, M., Fereshtehnejad, S.-M., Foreman, K.J., Gankpe, F.G., Gupta, R., Hamidi, S., Hankey, G.J., Hay, S.I., Hegazy, M.I., Hibstu, D.T., Kasaeian, A., Khader, Y., Khalil, I., Khang, Y.-H., Kim, Y.J., Kokubo, Y., Logroscino, G., Massano, J., Mohamed Ibrahim, N., Mohammed, M.A., Mohammadi, A., Moradi-Lakeh, M., Naghavi, M., Nguyen, B.T., Nirayo, Y.L., Ogbo, F.A., Owolabi, M.O., Pereira, D.M., Postma, M.J., Qorbani, M., Rahman, M.A., Roba, K.T., Safari, H., Safiri, S., Satpathy, M., Sawhney, M., Shafieesabet, A., Shiferaw, M.S., Smith, M., Szoeke, C.E.I., Tabarés-Seisdedos, R., Truong, N.T., Ukwaja, K.N., Venketasubramanian, N., Villafaina, S., weldegwergs, K. gidey, Westerman, R., Wijeratne, T., Winkler, A.S., Xuan, B.T., Yonemoto, N., Feigin, V.L., Vos, T., Murray, C.J.L.: Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 17, 939–953 (2018). https://doi.org/10.1016/S1474-4422(18)30295-3.
16. McKay, J.L., Lang, K.C., Bong, S.M., Hackney, M.E., Factor, S.A., Ting, L.H.: Abnormal center of mass feedback responses during balance: A potential biomarker of falls in Parkinson's disease. PLOS ONE. 16, e0252119 (2021). https://doi.org/10.1371/journal.pone.0252119.
17. Porter, M.F.: An algorithm for suffix stripping. Program. 14, 130–137 (1980). https://doi.org/10.1108/eb046814.
18. Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., Schneider, N., Smith, N.A.: Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters. Presented at the Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies June (2013).
19. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V.: RoBERTa: A Robustly Optimized BERT Pretraining Approach.
20. Harris, D.M., Duckham, R.L., Daly, R.M., Abbott, G., Johnson, L., Rantalainen, T., Teo, W.-P.: Development of a Parkinson's disease specific falls questionnaire. BMC Geriatr. 21, 614 (2021). https://doi.org/10.1186/s12877-021-02555-6.
21. Allen, N.E., Canning, C.G., Almeida, L.R.S., Bloem, B.R., Keus, S.H., Löfgren, N., Nieuwboer, A., Verheyden, G.S., Yamato, T.P., Sherrington, C.: Interventions for preventing falls in Parkinson's disease. Cochrane Database Syst. Rev. 2022, (2022). https://doi.org/10.1002/14651858.CD011574.pub2.
22. Duckham, R.L., Procter-Gray, E., Hannan, M.T., Leveille, S.G., Lipsitz, L.A., Li, W.: Sex differences in circumstances and consequences of outdoor and indoor falls in older adults in the MOBILIZE Boston cohort study. BMC Geriatr. 13, 133 (2013). https://doi.org/10.1186/1471-2318-13-133.