Trajectories in Physical Performance and Fall Prediction in Older Adults: A longitudinal population-based study

Kevin A Kerber, MD[1], Ran Bi, MS[2], Lesli E. Skolarus, MD[2], James F. Burke, MD [1]

[1] Department of Neurology, Ohio State University, Columbus OH
[2] Department of Neurology, University of Michigan, Ann Arbor MI

Funding: This work was supported by grant R01 MD008879 from the National Institute on Minority Health and Health Disparities.

Manuscript Word Count: 3500
Abstract Word Count: 293
Tables: 2; Figures 2

Corresponding Author:
Kevin A. Kerber, MD
Ohio State University
Department of Neurology
614.366.9092
kevin.kerber@osumc.edu

Key Points

- Among older adults with annual evaluations, we found that predicting time to ≥2 falls did not improve after adding data about the trajectory of physical performance to a model that used baseline physical performance and established non-performance-based information.
- These results did not address whether repeated physical performance measures are useful for other purposes.

Why does this paper matter?

The assessment of risk of falling is important and a required part of Annual Wellness Visits. Clinicians should use the most accurate and efficient methods to determine fall risk.
Since the trajectory of performance did not meaningfully improve falls prediction, we did not identify value to repeating this type of physical performance evaluation at each annual visit.

Abstract

Background. A physical performance evaluation can inform fall risk in older people, however the predictiveness of a one-time assessment is limited. The trajectory of physical performance over time has not been well characterized and might improve fall prediction. We aimed to characterize trajectories in physical performance and determine if fall prediction improves using trajectories of performance.

Methods: This was a cohort design using data from the National Health and Aging Trends Study. Physical performance was measured by the Short Physical Performance Battery (SPPB) with scores ranging from 0 (worst) to 12 (best). The trajectory of SPPB was categorized using latent class modelling and slope-based multilevel linear regression. We used Cox proportional hazards models with an outcome of time to ≥2 falls from annual self-report to assess predictiveness after adding SPPB trajectories to models of baseline SPPB and established non-physical-performance-based variables.

Results. The sample was 5,969 community-dwelling Medicare beneficiaries age ≥65 years. The median number of annual SPPB evaluations was 4(IQR,3-7). Mean baseline SPPB was 9.2(SD,3.0). The latent class model defined SPPB trajectories over a range of two to nineteen categories. The mean slope from the slope-based model was -0.01 SPPB points/year (SD,0.14). Discrimination of the baseline SPPB model to predict time to ≥2 falls was fair (Harrell's C,0.65) and increased after adding the non-performance-based predictors (Harrell's C,0.70). Discrimination slightly improved with the SPPB trajectory category variable that had the best fit

(Harrell's C,0.71) but did not improve with the SPPB linear slope. Calibration with and without the trajectory categories was similar.

Conclusions. We found that the trajectory of physical performance did not meaningfully improve upon fall prediction from a baseline physical performance assessment and established non-performance-based information. These results do not support longitudinal SPPB assessments for fall prediction.

Keywords. Falls, Physical performance, trajectories, older adults.

Introduction

Falls are the leading cause of fatal and nonfatal injuries among adults aged ≥65 years old.[1] A priority in healthcare is therefore to identify individuals at increased risk for falls and initiate fall prevention strategies.[2, 3] A bedside physical performance assessment – walking speed, standing balance, and chair rise – has been shown to be associated with falls and a predictor of future falls.[4-6] The evaluation is recommended by the Centers for Disease Control and Prevention as part of the process to assess fall risk in community dwelling older adults.[7] However, the discriminatory performance of a single assessment with falls was only fair (c-statistic,0.65-0.68).[4, 5]

Physical performance can change over time, and at different rates over time, based on a variety of factors such as aging, co-morbidities, injuries, and therapies. As a result, longitudinal assessments of physical performance and the associated trajectories in performance may improve the ability to predict future falls. A few previous studies of older adults have applied latent class modeling to describe the trajectory of physical performance.[8-10] These studies all identified three trajectories which were generally characterized as either a good baseline performance with minimal decline over time, an intermediate-good baseline performance with a mild-moderate decline, or an intermediate-low baseline performance with a substantial decline. A limitation of the latent class modeling, however, is that the results often lack much granularity since only a small number of trajectory categories are identified. The prior studies were also limited by having only 3-4 performance assessments available to define the trajectories. Finally, the prior studies did not evaluate whether the trajectory data improved fall prediction.

In this study, we sought to describe and define trajectories in physical performance using up to 8 years of data from a national sample of older adults in the National Health and Aging Trends Study (NHATS). We aimed to compare a latent class modeling approach to define trajectory categories with an approach that uses multi-level linear regression to determine a more granular slope at the individual level. To estimate the potential clinical utility of trajectories in physical performance, we then evaluated the marginal accuracy of predicting time to ≥2 falls by adding the trajectory data to models of baseline SPPB and established non-performance-based fall predictors (e.g., self-reported fear of falling, problems with balance, and use of an assistive device). We hypothesized that trajectory data would improve predictiveness and that slope trajectories would meaningfully outperform latent class trajectory categories.

Methods

Study Design & Data Sources

The design was a cohort study using data obtained prospectively from 2011-2018 in the National Health and Aging Trends Study (NHATS). NHATS is an annual in-home, longitudinal, nationally representative survey of community-dwelling Medicare beneficiaries 65 years and older drawn from the Medicare enrollment database.[11] NHATS oversamples individuals who self-report as black non-Hispanic and the oldest old (≥90 years). Trained staff perform annual in-person data collection from participants including self-reported measures and cognitive and physical assessments. Detailed methods of the NHATS have been published previously.[12] This study was approved by the University of Michigan Institutional Review Board.

Study Sample

The inclusion criterion was NHATS participants with two or more years of physical performance assessments. Participants were excluded from physical performance assessments when residing in a nursing home or other supportive living environments.[11] All other participants were screened to determine their eligibility for the physical performance assessment.[13] Because the study aimed to assess the marginal accuracy of predicting time to ≥2 falls by adding the trajectory data to baseline models, we also excluded participants who reported ≥2 falls at the baseline SPPB evaluation or had missing baseline covariates.

Physical Performance

NHATS physical performance measures are centered on the Short Physical Performance Battery (SPPB).[13-15] NHATS scores the SPPB from 0 (worst performance) to 12 (best performance) based on three activities: balance stand, walking speed, and repeated chair stand. Detailed protocols are available through NHATS (www.nhats.org).[13] The balance stand test was a progressive evaluation of the time the participant could hold standing positions with feet side-by-side, in semi-tandem, in full-tandem, and then standing on one leg with eyes open and then eyes closed. The time was limited to up to 10 seconds for the two-leg stands or 30 seconds for the one-leg stands, or when the participant stepped out of position or grabbed the interviewer's arm. For walking speed, the time to walk 3 meters was recorded and converted to meters per second. For repeated chair stands, participants were instructed to stand up from the chair, with arms folded across their chest, as quickly as possible in one minute up to five times. The test was stopped if participants used their arms to aid in standing up, did not complete five rises within one minute, or displayed a behavior that raised a safety concern (e.g., shortness of breath). Each of the activities was scored 0 to 4 and summed to obtain the SPPB score. Scores from 1 to 4 on the activities were based on quartiles of the weighted distribution for non-missing, non-zero values. Participants received a score of zero for an activity if they were not eligible for the task (receipt of help, use of assisted device, or had surgery on both hips within 3 months), did not attempt the task, or safety concerns were identified. If the participant, proxy, or interviewer felt the task was not safe to try, the participant also received a score of zero.

Falls

Information about falls was obtained in NHATS by self-report. Falls were defined as any fall, slip, or trip in which the participant lost their balance and landed on the floor or ground or at a lower level.[11] Participants were first asked, "In the last 12 months, have you fallen down?" Individuals who responded yes were then asked, "In the last 12 months, have you fallen down more than once?"

Established Fall Predictor Variables

Covariate predictors of falls were selected to match fall risk screening items from the American Geriatric Society (AGS) and the Centers for Disease Control (CDC) STEADI initiative.[2, 3] These variables included questions about fear of falling, problems with balance, use of cane or walker, hearing difficulty, vision difficulty, depression, environmental hazards (i.e., floor needs repair, other tripping hazards), and cognitive status (Table S1). Cognitive status is classified in NHATS as no dementia, possible dementia, or probable dementia.[11] AGS and CDC falls risk screening questions that were not available in NHATS were the following: need to push to stand up from a chair, trouble stepping off a curb, rushing to the toilet, loss of feeling in the feet, and taking medicine that causes lightheadedness.

Demographic Characteristics

Age, sex, self-identified race/ ethnicity, education, and marital status, as reported in the baseline interview, were included. Race/ethnicity was categorized into White non-Hispanic, Black non-Hispanic, or Hispanic. Additional non-Hispanic groups including American Indian, Asian, and Native Hawaiian/Pacific Islander were reported as other due to small sample sizes.

Statistical Analysis

We used descriptive statistics to summarize the demographic and clinical characteristics of the sample including survey weights to account for the complex design, overall and by category of baseline SPPB performance (Poor, 0-6; Intermediate, 7-9; Good, 10-12).[8] For all variables, a 'don't know' response was recoded as either a no response or the lowest ranked category (e.g., education). To determine trajectories of physical performance, we used two different methods: a latent-class modelling approach, and a slope-based modelling approach. The advantage of the latent-class model is that it has been relatively widely used,[8-10, 16] but the potential disadvantage is it only typically identifies a small number of trajectory categories. The slope-based modelling approach, on the other hand, calculates granular slopes at the individual level, but cannot account for non-linearities in change over time.

To estimate a latent-class model of SPPB, we used Stata command “traj”.[17, 18] This is a specialized form of finite mixture modelling and is designed to identify latent classes of individuals following similar progression of a variable of interest over time. The approach uses a multinomial modeling strategy and maximum likelihood for the estimation of the model parameters. We used a censored normal distribution and linear polynomial types for each group trajectory. Cases were censored at ≥2 falls. The time variable was year of the SPPB evaluation ranging from 2-8. For every participant, the analysis calculates the posterior probabilities for each trajectory. We assigned participants to the trajectory with the highest probability. Because the number of latent classes is unknown, we estimated models across a range of potentially distinct trajectories starting at two categories and continuing until reaching the maximum Bayesian Information Criterion as calculated by Nagin.[19] To evaluate the models, we also determined the Akaike information criterion, entropy, and the lowest average of the posterior probability of the group memberships.[20, 21] For the slope-based model, we used multilevel mixed effects linear regression to calculate the slope of the SPPB per individual for every year after the second year of SPPB performance. The slopes were updated as each additional year of SPPB was added. The final slope per individual was taken from their last year of SPPB evaluation.

We then used a series of Cox proportional hazard models to examine the independent association of the performance trajectories and time to ≥2 falls in the past 12 months before and after adjusting for the falls risk covariates. For all models, predictor variables were time lagged with the falls outcome because the fall outcome specifically queries events in the last 12 months. The models that examined the association of the slope-based trajectory with time to ≥2 falls used the slope from the last available SPPB performance year for each individual. Cases were censored for death and for receiving a score of zero on the SPPB evaluation (as assigned based on being ineligible for the task, not attempting the task, or safety concerns). The censoring for a score of zero was used as an attempt to address previously reported calibration problems in the highest predicted risk individuals which may be due to reduced mobility or the initiation of fall risk mitigation interventions.[22, 23] In the primary models, the AGS/CDC falls predictor variables were from the baseline assessment. Cox model discrimination was assessed using two measures: the Harrell's C coefficient, which depends on the unknown censoring distribution, and the Gönen and Heller's K concordance coefficient,[24] which does not. Model fit was assessed using the Bayesian Information Criterion.[25] Cox model calibration was evaluated visually by plotting predicted event probability by observed event probability.[26] We also calculated the median (IQR) predicted probability of ≥2 falls at each year by multiplying the mean baseline cumulative hazard function of each year by the exponentiated linear predictor. Three sensitivity analyses were performed. First, we performed an analysis with the covariates and the SPPB variable as time varying. Second, we evaluated the possibility of competing events biasing the results by changing the outcome to time to ≥2 falls, move to a nursing home or supportive living environment, or death. Third, to evaluate if cognitive status modifies the effect of SPPB on time to ≥2 falls, we performed a post hoc analysis adding an interaction term of SPPB trajectories and cognitive status. The complex survey design was accounted for in all analyses by applying survey weights. All analyses were performed using Stata (version 17; Stata Inc., College Station, TX).

Results

Study sample.

From 2011-2018, there were 12,427 adults aged 65 years or older in NHATS. The final study sample was 5,969 after excluding participants who did not have SPPB data due to living in a nursing home or other supportive living environment (N=869), who did not have at least 2 SPPB evaluations (N=2981), who reported ≥2 falls in the last 12 months at the baseline SPPB assessment year (N=2515), or who had missing covariates (N=92)(Figure 1).

The baseline characteristics of the study sample are presented in Table 1 overall and by SPPB performance category. In the cohort, 1968 participants (20%) were aged 80 years or older, 3473(56%) were female, and 4186(82%) were white. At the baseline evaluation, the median SPPB performance was 10 (IQR,8-12;range 0-12; mean,9.2 (SD, 3.0). Individuals with poor baseline SPPB performance were substantially older than those in the other baseline performance categories (Table 1). The baseline poor SPPB performers were also more frequently female, Black non-Hispanic, and had a higher frequency of falls in the baseline year, concerns about falling, self-reported problems with balance, use of cane or walker, problems with hearing, problems with vision, depression, possible/probable dementia, and environmental hazards (Table 1).

Longitudinal Characteristics & Trajectory of Physical Performance

The median number of annual SPPB evaluations was 4 (IQR 3-7;range 2-8). The SPPB score changed from the baseline to the final assessment by a median of -1 point (IQR, -2, 0) on the 12-point scale. The baseline SPPB score was strongly correlated with the subsequent annual SPPB scores although there was a gradual decrease in the correlation over time (r=0.78 from year 1-2 to 0.62 from year 1-7) (Table S2).

The results of the latent-class models are shown in Table S3 and Figure S1. The model with 17 trajectory categories had the best fit based on the Bayesian Information Criterion. However, the lowest posterior probability of group assignment, entropy, and size of the smallest group all declined as the number of trajectory categories increased. In the slope-based model, the median slope at the final year was -0.01 points per assessment (IQR, -0.13,0.07; range, -1.04,0.80; mean, -0.01,SD, 0.14).

Falls & Time to ≥2 Falls in the Past 12 Months

Over the study period, 22% of the sample reported ≥2 falls in the prior 12-months at one of the assessments, which varied from 18% for the baseline good performance category to 33% for the baseline poor performing category (Table 1).

The results of the Cox proportional hazards models to predict time to ≥2 falls are displayed in Table 2. The discrimination of the model that included only baseline SPPB performance was fair (Model 1: Harrell's C,0.65). Discrimination moderately improved when the baseline non-performance predictors were added to the model (Harrell's C,0.70). After adding the SPPB trajectory categories to the model, small additional gains in discrimination were noted (Model 3, Harrell's C 0.70-0.73; Tables 2 & S4). The Cox model with the best fit based on the Bayesian Information Criterion had 9-trajectory categories (Tables 2 & S4). Adding the slope trajectory to the baseline SPPB and non-performance predictors model did not increase discrimination (Model

4, Harrell's C, 0.70). Figure 2 displays the calibration of Model 2 and Model 3 at year 4. Both models were generally well calibrated in the range of ~20% to ~50% predicted probability of ≥2 falls. However, both models overpredicted the probability of ≥2 falls in individuals with >50% predicted probability (Figure 2). The predicted cumulative probability of ≥2 falls at each year was similar for Model 2 and Model 3 (Table S5). In the sensitivity analysis with time varying covariates including the SPPB, there was an increase of the associations of some of the non-performance variables (fell once in last 12 months, fear of falling, depressed, use of cane or walker, hearing problems, probable dementia) with time to ≥2 falls (Table S6). However, there was attenuation of the associations of the SPPB variable and the SPPB trajectory categories with time to ≥2 falls. The addition of the interaction term of dementia status with trajectory category led to little change in model discrimination (Harrell's C,0.71). In the sensitivity analysis that changed the outcome to time to ≥2 falls, facility placement, or death, the associations of the individual predictors and the overall model discrimination improved (Harrell's C,0.76).

Discussion

In this cohort study of nearly 6000 community-dwelling older adults with up to 8 years of annual assessments, we identified two key findings. First, we found that a latent class model that defined SPPB trajectories had higher discriminatory performance for predicting time to ≥2 falls than the more granular linear slope-based modeling approach. Second, we found that the SPPB trajectory categories did not meaningfully improve the discriminatory performance of time to ≥2 falls compared with the prediction that only used the baseline SPPB and non-performance-based variables. Overall, these findings indicate that a single baseline assessment of physical performance is likely sufficient for informing fall risk prediction with little additional value from subsequent physical performance assessments.

This study raises important questions about the value of physical performance evaluations, such as the SPPB, at annual visits for older people. The CDC STEADI initiative recommends evaluating gait, strength, and balance annually for the purpose of assessing fall risk.[3] The Centers for Medicare and Medicare Services (CMS) requires a falls risk assessment as part of an annual wellness visit but states that this can be done by either observing functional performance or using established screening questions.[27] It takes time and space to assess performance of gait, balance, and strength particularly when a formal scale, such as the SPPB, is used. Since the trajectory of SPPB performance did not meaningfully improve falls prediction, our findings therefore did not identify value to repeating this type of physical performance evaluation at each annual visit. The limited value of repeating physical performance was further highlighted by the time-varying model that showed there was a shift toward greater relative effects of non-performance-based variables compared to SPPB for fall prediction.

To our knowledge, this is the first study of older adults to characterize physical performance trajectories using multiple methods and to evaluate the marginal value of the trajectories on predicting time to ≥2 falls. Prior studies have characterized trajectories in physical performance using latent models but did not evaluate alternative strategies to define trajectories or the predictiveness of the trajectories with future falls.[8-10]

Our findings about the number of trajectories in SPPB performance differed from prior studies.[8-10] Prior studies concluded that three trajectories had the best fit of the data, whereas we found a better fit with a larger number of trajectories. This difference likely stems from our much larger sample size (prior studies with 604-1400 participants compared with our 5969) which facilitated the identification of additional trajectories.[8-10]

Our findings did not support our hypotheses that a more granular trajectory from the slope-based approach would be superior to the latent class modeling approach or that the trajectory data would improve the accuracy of fall prediction from models that used only baseline data. While the slope-based modelling approach necessarily results in a more granular characterization of the SPPB trajectory at the individual level, it did not outperform the trajectory category because the slopes had a narrow range of values across the sample, in part due the constraints of a linear model. In addition, an advantage of the latent class modeling approach is that that it relaxes the linearity assumption implicit in a continuous baseline SPPB score as a predictor of falls risk.[28] The main reason that the trajectory categories did not meaningfully improve fall prediction was likely that SPPB performance scores over the study period remained relatively stable (i.e., only changing 0-2 points on the 12-point scale) for most individuals. Consistent with this finding, the baseline performance was highly correlated with the performance in subsequent years. Another likely reason that the trajectory data did not meaningfully improve fall prediction was that ≥2 falls over a 12-month period was common (~20%) even in the individuals with good baseline performance. Lastly, the trajectory categories did not resolve the poor calibration at the highest probability of falls.

This study has several important strengths. First, the NHATS is a prospective study that includes annual SPPB evaluations and collects all the study's data with structured processes and trained staff.[11] Our study also had 8 annual assessments. The NHATS therefore had an optimal design for our goal of characterizing SPPB trajectories and evaluating their potential value in fall prediction for older adults. An additional strength of our analysis to predict time to ≥2 falls was that we censored individuals for safety concerns about SPPB performance since safety concerns should already indicate substantial fall risk and justify fall prevention initiatives – conditions that create the potential for poor calibration in those with high predicted fall risk. The lack of similar censoring in prior studies was a potential reason for their poor calibration.[22, 23] However, our censoring did not eliminate the poor calibration at the high end of predicted probability.

This study also had important limitations. The outcome variable of ≥2 falls in the past 12 months was based on self-report. In addition, we did not have data on falls with injury. Due to available data from NHATS, our models for fall prediction were not able to include all of the CDC non-performance-based fall predictors. Predictors not included were questions about loss of feeling in the feet, taking medicines that cause lightheadedness, taking medicines that increase fall risk, or the presence of orthostatic hypotension. We also could not adjust for individuals who had undergone fall prevention interventions. We excluded participants with missing data in the model variables. Since the frequency of participants with missing data was small (~1.5%; 92/6961), it is unlikely that this biased the results. It is possible that participants had competing events that were not accounted for and could have biased the results. We explored the possibility of two competing events – death or facility placement – as a sensitivity analysis but the results were similar to our primary model. Other competing events that we did not consider may exist.

Conclusions

In conclusion, a slope-based SPPB trajectory model did not outperform the latent class model to determine trajectory categories for discriminating time to ≥2 falls. In addition, the SPPB trajectory data did not meaningfully improve the fall prediction that used baseline only information because discrimination only marginally improved but calibration did not. These findings do not support the use of annual SPPB evaluations for a falls risk assessment in older adults.

Acknowledgments: This work was supported by grant R01 MD008879 from the National Institute on Minority Health and Health Disparities.

Sponsor's Role: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Conflict of Interest: The authors have no conflicts.

Author Contributions: Study concept and design (KK, JB, LS), acquisition of subjects and/or data (RB), analysis and interpretation of data (RB, KK, JB, LS), and preparation of manuscript (KK, RB, JB, LS).

References

1. Bergen G, Stevens MR, Burns ER. Falls and Fall Injuries Among Adults Aged >/=65 Years - United States, 2014. MMWR Morb Mortal Wkly Rep 2016;65:993-998.

2. 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 Am Geriatr Soc 2011;59:148-157.
3. Centers for Disease Control and Prevention. STEADI-Older Adult Fall Prevention. https://www.cdc.gov/steadi/index.html.
4. Palumbo P, Palmerini L, Bandinelli S, Chiari L. Fall Risk Assessment Tools for Elderly Living in the Community: Can We Do Better? PLoS One 2015;10:e0146247.
5. Lauretani F, Ticinesi A, Gionti L, et al. Short-Physical Performance Battery (SPPB) score is associated with falls in older outpatients. Aging Clin Exp Res 2019;31:1435-1442.
6. Veronese N, Bolzetta F, Toffanello ED, et al. Association between Short Physical Performance Battery and falls in older people: the Progetto Veneto Anziani Study. Rejuvenation Res 2014;17:276-284.
7. STEADI: Older Adult Fall Prevention. Centers for Diease Control and Prevention. https://www.cdc.gov/steadi/index.html.
8. Mutambudzi M, Chen NW, Howrey B, Garcia MA, Markides KS. Physical Performance Trajectories and Mortality Among Older Mexican Americans. J Gerontol A Biol Sci Med Sci 2019;74:233-239.
9. Ahmed T, French SD, Belanger E, Guerra RO, Zunzunegui MV, Auais M. Gender Differences in Trajectories of Lower Extremity Function in Older Adults: Findings From the International Mobility in Aging Study. J Am Med Dir Assoc 2019;20:1199-1205 e1194.
10. Charles A, Detilleux J, Buckinx F, Reginster JY, Gruslin B, Bruyere O. Physical performance trajectories and mortality among nursing home residents: results of the SENIOR cohort. Age Ageing 2020;49:800-806.
11. Kasper JD, Freedman V. National Health and Aging Trends Study User Guide: Rounds 1-10 Final Release. 2021.
12. Freedman VA, Kasper JD, Cornman JC, et al. Validation of new measures of disability and functioning in the National Health and Aging Trends Study. J Gerontol A Biol Sci Med Sci 2011;66:1013-1021.
13. Kasper JD, Freedman VA, Niefeld MR. Construction of performance-based summary measures of physical capacity in the National Health and Aging Trends Study. NHATS Technical Paper #4. . 2012.
14. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49:M85-94.
15. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med 1995;332:556-561.
16. Wolf DA, Freedman VA, Ondrich JI, Seplaki CL, Spillman BC. Disability Trajectories at the End of Life: A "Countdown" Model. J Gerontol B Psychol Sci Soc Sci 2015;70:745-752.
17. Nagin DS, Jones BL, Passos VL, Tremblay RE. Group-based multi-trajectory modeling. Stat Methods Med Res 2018;27:2015-2023.
18. Jones BL, Nagin DS. A note on a Stata plugin for estimating group-based trajectory models. Sociological Methods & Research 2013:608-613.
19. Nagin D. Group-based modeling of development. : Harvard University Press, 2009.
20. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol 2010;6:109-138.
21. van de Schoot R, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK. The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. Structural Equation Modeling: A Multidisciplinary Journal 2017;24:451-467.

22. Cattelani L, Palumbo P, Palmerini L, et al. FRAT-up, a Web-based fall-risk assessment tool for elderly people living in the community. J Med Internet Res 2015;17:e41.
23. Palumbo P, Klenk J, Cattelani L, et al. Predictive Performance of a Fall Risk Assessment Tool for Community-Dwelling Older People (FRAT-up) in 4 European Cohorts. J Am Med Dir Assoc 2016;17:1106-1113.
24. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. The Stata Journal 2010;10:339-358.
25. Schwarz G. Estimating the dimension of a model. Annals of Statistics 1978;6:461-464.
26. Royston P. Tools for checking calibration of a Cox model in external validation: Approach based on individual event probabilities. The Stata Journal 2014;14:738-755.
27. National Archives and Record Administration Code of Federal Regulations https://wwwecfrgov/current/title-42/chapter-IV/subchapter-B/part-410/subpart-B/section-41015.
28. Quach L, Galica AM, Jones RN, et al. The nonlinear relationship between gait speed and falls: the Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly of Boston Study. Journal of the American Geriatrics Society 2011;59:1069-1073.

Table 1. Characteristics of the study sample.*

	Full Sample N, weighted proportion unless otherwise specified	Poor SPPB Baseline^ N, weighted proportion unless otherwise specified	Intermediate SPPB Baseline^ N, weighted proportion unless otherwise specified	Good SPPB Baseline^ N, weighted proportion unless otherwise specified
N	5969 (1.0)	1447 (0.17)	1526 (0.23)	2996 (0.61)
Age (years)				
65 - 69	1393 (0.37)	124 (0.15)	265 (0.28)	1004 (0.46)
70 - 74	1363 (0.25)	215 (0.19)	319 (0.23)	829 (0.27)
75 - 79	1245 (0.18)	271 (0.20)	372 (0.22)	602 (0.16)
80 - 84	1077 (0.12)	367 (0.22)	330 (0.16)	380 (0.07)
85 - 89	573 (0.06)	271 (0.16)	171 (0.08)	131 (0.02)
>= 90	318 (0.02)	199 (0.09)	69 (0.02)	50 (0.01)
Female	3473 (0.56)	1027 (0.69)	931 (0.60)	1515 (0.51)
Race/Ethnicity				
White, non-Hispanic	4186 (0.82)	803 (0.70)	996 (0.76)	2387 (0.88)
Black, non-Hispanic	1266 (0.08)	476 (0.15)	379 (0.10)	411 (0.05)
Other	170 (0.03)	48 (0.04)	41 (0.04)	81 (0.03)
Hispanic	339 (0.07)	117 (0.10)	110 (0.10)	112 (0.04)
Missing	8 (0.00)	3 (0.00)	0 (0.00)	5 (0.00)
Education				
< High school	1281 (0.16)	538 (0.32)	383 (0.22)	360 (0.09)
High school	1587 (0.25)	401 (0.29)	455 (0.29)	731 (0.23)
> High school	3097 (0.59)	507 (0.39)	688 (0.49)	1902 (0.68)
Missing	4 (0.00)	1 (0.00)	0 (0.00)	3 (0.00)
Married	3018 (0.58)	431 (0.35)	739 (0.53)	1848 (0.66)
Missing	5 (0.00)	3 (0.00)	0 (0.00)	2 (0.00)
Health Status				
Excellent	910 (0.18)	79 (0.06)	153 (0.10)	678 (0.24)
Very good	1891 (0.35)	268 (0.20)	434 (0.29)	1189 (0.41)
Good	1951 (0.30)	496 (0.34)	599 (0.39)	856 (0.26)
Fair	977 (0.14)	445 (0.30)	284 (0.19)	248 (0.07)
Poor	240 (0.03)	159 (0.10)	56 (0.03)	25 (0.01)
Fall in baseline wave	1400 (0.23)	480 (0.35)	364 (0.26)	556 (0.19)

Worry about falling down	1333 (0.20)	599 (0.45)	363 (0.25)	371 (0.12)
Often feel depressed	319 (0.05)	143 (0.11)	93 (0.05)	83 (0.03)
Have problems with balance	1367 (0.19)	691 (0.48)	360 (0.24)	316 (0.10)
Use a cane or walker	1277 (0.16)	872 (0.57)	270 (0.17)	135 (0.04)
Hearing problems	139 (0.02)	53 (0.03)	37 (0.03)	49 (0.01)
Vision problems	315 (0.04)	169 (0.11)	70 (0.05)	76 (0.02)
Cognitive status				
No dementia	4949 (0.88)	902 (0.66)	1284 (0.87)	2763 (0.94)
Possible dementia	649 (0.08)	287 (0.18)	179 (0.10)	183 (0.05)
Probable dementia	371 (0.04)	258 (0.16)	63 (0.03)	50 (0.01)
Environmental Hazard				
No	5081 (0.86)	1144 (0.80)	1302 (0.85)	2635 (0.89)
Inapplicable	286 (0.05)	80 (0.06)	58 (0.04)	148 (0.05)
Yes	602 (0.08)	223 (0.14)	166 (0.11)	213 (0.06)
Baseline SPPB, median (IQR)	10 (8-12)	4 (1-5)	8 (8-9)	11 (11-12)
Baseline SPPB, mean (SD)	9.23 (2.96)	3.36 (2.51)	8.19 (0.80)	11.22 (0.68)
Final SPPB, mean (SD)	8.40 (3.57)	3.50 (3.87)	7.41 (3.26)	10.11 (2.17)
≥2 falls over observation period	1463 (0.22)	457 (0.33)	408 (0.26)	598 (0.18)

SPPB= Short Physical Performance Battery; SD= standard deviation; IQR = interquartile range
*All statistics calculated using survey weights.
^Baseline categories of SPPB defined as Poor (0-6), Intermediate (7-9), and Good (10-12).

Table 2. Cox proportional hazard models to predict time to ≥2 falls.

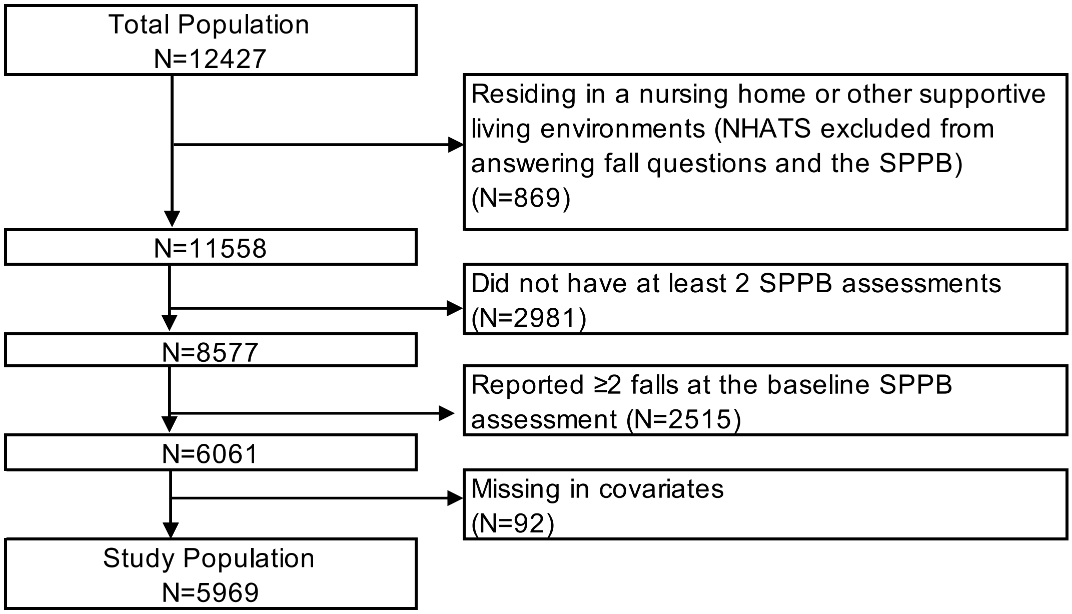
	Model 1: Baseline SPPB only HR (95% CI) (N=5969)	Model 2: Baseline SPPB + Non-Performance based variables HR (95% CI) (N=5969)	Model 3:* Baseline SPPB + Non-Performance based variables + SPPB trajectories HR (95% CI) (N=5969)	Model 4: Baseline SPPB + Non-Performance based variables + Trajectory Slope HR (95% CI) (N=5969)
SPPB Baseline	0.87 (0.85-0.88)	0.94 (0.92-0.96)	0.96 (0.93-1)	0.94 (0.92-0.96)
Trajectory Group				
1			Ref	
2			0.27 (0.18-0.4)	
3			0.08 (0.03-0.22)	
4			0.17 (0.09-0.3)	
5			0.28 (0.18-0.44)	
6			0.20 (0.13-0.31)	
7			0.26 (0.16-0.43)	
8			0.20 (0.12-0.35)	
9			0.10 (0.05-0.2)	
Trajectory Slope				1.01 (0.70-1.46)
Covariates				
Fell once last 12m		1.72 (1.54-1.92)	1.71 (1.53-1.92)	1.72 (1.54-1.92)
Fear of falling		1.24 (1.06-1.44)	1.22 (1.05-1.43)	1.24 (1.06-1.45)
Depressed		1.06 (0.78-1.44)	1.12 (0.86-1.47)	1.06 (0.78-1.44)
Problems with balance		1.68 (1.39-2.04)	1.64 (1.35-1.98)	1.68 (1.39-2.04)
Use of cane or walker		1.33 (1.11-1.60)	1.31 (1.09-1.57)	1.33 (1.11-1.60)

Hearing problems		1.11 (0.79-1.56)	1.21 (0.88-1.67)	1.11 (0.79-1.56)
Vision problems		1.12 (0.87-1.44)	1.04 (0.81-1.33)	1.12 (0.87-1.44)
Dementia				
Possible		1.32 (1.10-1.57)	1.35 (1.14-1.6)	1.32 (1.10-1.58)
Probable		1.65 (1.26-2.17)	1.48 (1.13-1.94)	1.65 (1.26-2.16)
Environmental Hazard				
Inapplicable		0.87 (0.65-1.15)	0.87 (0.66-1.16)	0.87 (0.65-1.15)
1 or more		1.17 (1.00-1.37)	1.19 (1.03-1.38)	1.17 (1.00-1.37)
*Harrell's C	0.6482	0.7001	0.7125	0.6999
*Gonen & Heller's K	0.6158	0.6474	0.6666	0.6474

HR = hazard ratio; SPPB = Short Physical Performance Battery
*Trajectory with 7 categories presented based the best fit Bayesian Information Criterion.

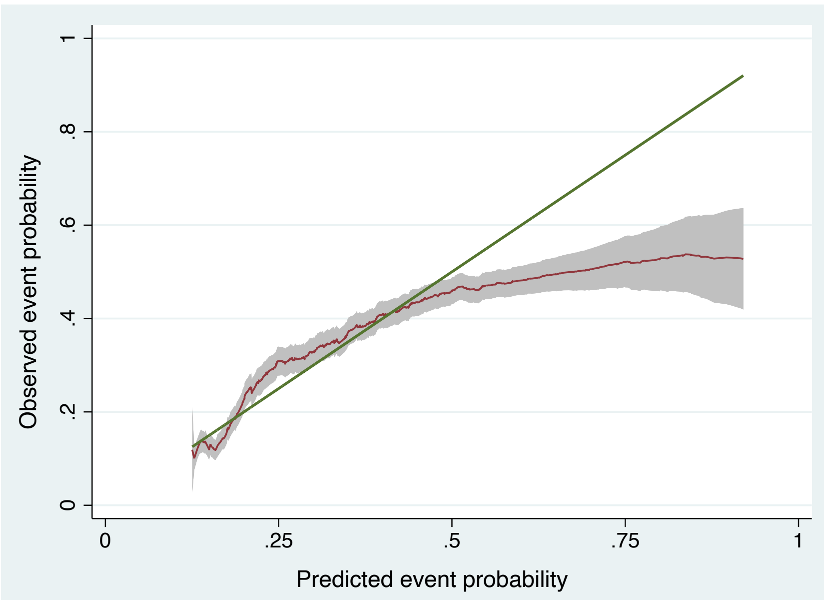
Figure 1. Flow diagram of study sample.

Figure 2. Calibration plots from Cox proportional hazard models. A) Calibration from model with baseline SPPB and non-performance variables. B) Calibration from model with baseline SPPB, SPPB trajectory categories, and non-performance variables. SPPB= Short Physical Performance Battery. The time point at which calibration was assessed was year 4 since the median number of annual SPPB evaluations was 4.

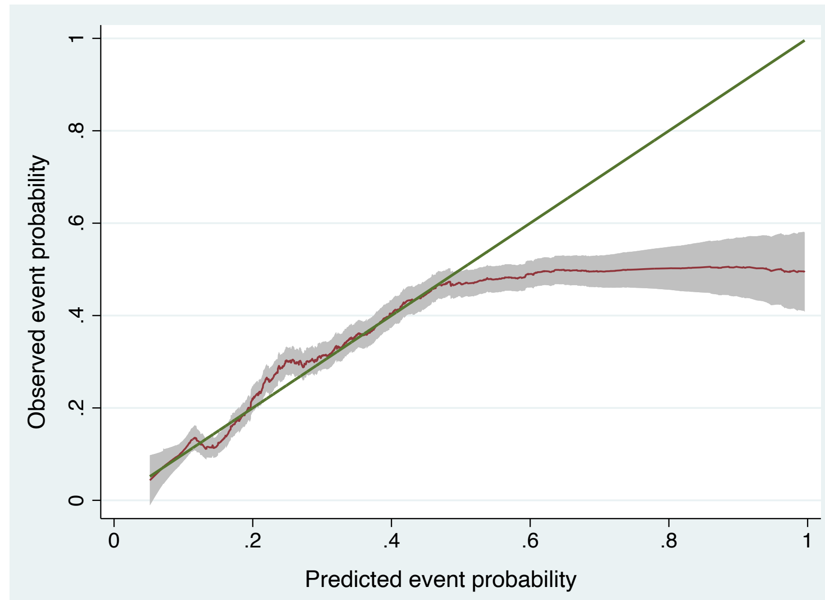


Total Population
N=12427
Residing in a nursing home or other supportive living environments (NHATS excluded from answering fall questions and the SPPB) (N=869)
N=11558
Did not have at least 2 SPPB assessments (N=2981)
N=8577
Reported ≥2 falls at the baseline SPPB assessment (N=2515)
N=6061
Missing in covariates (N=92)
Study Population
N=5969

A.)


Observed event probability
0
.2
.4
.6
.8
1
0
.25
.5
.75
1
Predicted event probability

B.)


Observed event probability
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
Predicted event probability