HHS Public Access

Author manuscript

J Am Med Dir Assoc. Author manuscript; available in PMC 2018 September 13.

Published in final edited form as:

J Am Med Dir Assoc. 2016 December 01; 17(12): 1106–1113. doi:10.1016/j.jamda.2016.07.015.

Predictive Performance of a Fall Risk Assessment Tool for Community-Dwelling Older People (FRAT-up) in 4 European Cohorts

Pierpaolo Palumbo, PhD[a,*], **Jochen Klenk, PhD**[b,c], **Luca Cattelani, PhD**[d], **Stefania Bandinelli, MD**[e], **Luigi Ferrucci, MD, PhD**[f], **Kilian Rapp, MD, MPH**[b], **Lorenzo Chiari, PhD**[a,g], and **Dietrich Rothenbacher, MD, MPH**[c]

[a]Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi"–DEI, University of Bologna, Bologna, Italy

[b]Department of Geriatrics and Clinic of Geriatric Rehabilitation, Robert-Bosch-Hospital, Stuttgart, Germany

[c]Institute of Epidemiology and Medical Biometry, Ulm University, Ulm, Germany

[d]Department of Computer Science and Engineering–DISI, University of Bologna, Bologna, Italy

[e]Local Health Unit Tuscany Center, Florence, Italy

[f]Clinical Research Branch, National Institute on Aging, Baltimore, MD

[g]Health Sciences and Technologies Interdepartmental Center for Industrial Research, University of Bologna, Bologna, Italy

Abstract

Background and objective: The fall risk assessment tool (FRAT-up) is a tool for predicting falls in community-dwelling older people based on a meta-analysis of fall risk factors. Based on the fall risk factor profile, this tool calculates the individual risk of falling over the next year. The objective of this study is to evaluate the performance of FRAT-up in predicting future falls in multiple cohorts.

Methods: Information about fall risk factors in 4 European cohorts of older people [Activity and Function in the Elderly (ActiFE), Germany; English Longitudinal Study of Aging (ELSA), England; Invecchiare nel Chianti (InCHIANTI), Italy; Irish Longitudinal Study on Aging (TILDA), Ireland] was used to calculate the FRAT-up risk score in individual participants. Information about falls that occurred after the assessment of the risk factors was collected from subsequent longitudinal follow-ups. We compared the performance of FRAT-up against those of other prediction models specifically fitted in each cohort by calculation of the area under the receiver operating characteristic curve (AUC).

[*]Address correspondence to Pierpaolo Palumbo, PhD, Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi"–DEI, Biolab, University of Bologna, Viale del Risorgimento, 2, I-40136 Bologna (BO), Italy. pierpaolo.palumbo@unibo.it (P. Palumbo).

Results: The AUC attained by FRAT-up is 0.562 [95% confidence interval (CI) 0.530–0.594] for ActiFE, 0.699 (95% CI 0.680–0.718) for ELSA, 0.636 (95% CI 0.594–0.681) for InCHIANTI, and 0.685 (95% CI 0.660–0.709) for TILDA. Mean FRAT-up AUC as estimated from meta-analysis is 0.646 (95% CI 0.584–0.708), with substantial heterogeneity between studies. In each cohort, FRAT-up discriminant ability is surpassed, at most, by the cohort-specific risk model fitted on that same cohort.

Conclusions: We conclude that FRAT-up is a valid approach to estimate risk of falls in populations of community-dwelling older people. However, further studies should be performed to better understand the reasons for the observed heterogeneity across studies and to refine a tool that performs homogeneously with higher accuracy measures across different populations.

Keywords

Older people; falls; FRAT-up; prediction model; validation

For many age-related health conditions or health-related threats, information about epidemiologic measures, such as incidence and prevalence, knowledge of the natural history, risk factors, or risk indicators, has allowed the development of condition-specific predicttion tools.[1–8] Such tools express the likelihood that an individual under assessment will experience the undesired condition of interest within a given time span. They are used in public health, medical research, and clinical practice for identification of high-risk persons who can be targeted for cost-effective preventive interventions.[9–11]

Falls are highly prevalent in older people. They are associated with increased morbidity and even mortality. Falls are a major cause of deterioration in quality of life because they can result in physical injuries (eg, fractures) and negative psychological attitudes, such as loss of self-efficacy. Fall prevention interventions can benefit from valid fall prediction tools.[12–14] Although many such tools have been proposed, only a few of them have been extensively validated and have been found to have only modest predictive accuracy.[6,15–21]

Recently, Cattelani et al[22] proposed a new prediction tool for falls in community-dwelling older people called (FRAT-up). It calculates the risk of falling for an individual, expressed as the probability of falling within the next 12 months. The tool is freely available online.[23] Its architecture can be outlined as the cascade of 2 building blocks. The first block receives some clinical variables of the person under observation, that the authors called “risk estimators”, and estimates the person’s exposure to a list of FRAT-up-defined fall risk factors. The second block uses this information about exposure to the risk factors and calculates the probability of falling. When applying FRAT-up on datasets of different studies, the first block acts as a “harmonization block,” which adapts to different risk estimators included in each dataset (ie, different clinical scales, medical instruments, or protocols) and converts this information into risk factor exposures (ie, whether the person has vision impairments, gait problems, etc). The second block remains unchanged across different datasets and can be considered as the “core block.” This architecture makes FRAT-up a flexible tool and allows it to be used across studies where different risk estimators were used to estimate the fall risk factors, which is the usual case.

All the parameters of the core block of FRAT-up were derived from the literature. In particular, the parameters that determine the contribution of each risk factor to the overall risk of falling were determined from the odds ratios obtained in the systematic review and meta-analysis by Deandrea et al.[24] Until now, FRAT-up has been evaluated only in the Invecchiare nel Chianti (InCHIANTI) cohort.[22] However, because the meta-analysis by Deandrea collated results from numerous epidemiologic studies, with risk factors assessed by different risk estimators, we hypothesize that FRAT-up is a suitable screening tool for different populations and can be adapted to different methods for risk factors assessment (ie, different risk estimators).

With the present study, we aim to further validate FRAT-up and verify this hypothesis, evaluating its predictive performance on 4 datasets from relevant European epidemiologic studies including community-dwelling older adults. The performance of a predictive model depends on the model itself but also on the cohort on which it is tested. To gain better insight on the robustness of FRAT-up performance across different datasets, we also aim to compare the predictive performance of FRAT-up with data-driven prediction models, each specifically fitted on 1 of the 4 cohorts.

Methods

The FRAT-up validation process is described in this article in compliance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis checklist for transparent reporting.[25,26] To achieve the objectives listed above, we used 4 datasets from cohort studies conducted in different European countries (Germany, England, Italy, and Ireland). The 4 datasets were initially harmonized to obtain estimates of risk exposure on a standard list of risk factors. The FRAT-up risk score was calculated and 4 cohort-specific prediction models were developed for comparison. All analyses were run with R version 3.0.2 (R Core Team, Vienna, Austria).[27]

Included Study Populations

The Activity and Function in the Elderly (ActiFE) in Ulm study is a population-based observational study on a cohort of community-dwelling older adults. Its principal aim is to investigate the relation-ship of physical activity, measured with body-worn accelerometers, with a number of health outcomes. The study design has been pre-viously described in detail.[28,29] Briefly, inclusion criteria were living in the area of greater Ulm or Neu-Ulm, located in the South of Germany; being 65 to 90 year old; not being institutionalized; being able to walk independently through their own room; not having serious difficulties in German language, and no severe deficits in cognition. Older age strata were oversampled to recruit an equal number of persons for each age group. At baseline (2009–2010), 1506 participants were assessed on a number of health parameters, including the fall risk estimators used in the present study. Successively, they were prospectively followed for 12 months to monitor the occurrence of falls using fall calendars as recommended by the Profane consortium.[30] We excluded 90 people (6%) on whom follow-up information about falls was missing.

The English Longitudinal Study of Aging (ELSA) is a panel study of a cohort that is representative of the population of noninstitutionalized men and women aged 50 years or older living in England. Its broad scope is to study aging in England in its health, economic, and social aspects.[31] In 2004–2005 (wave 2), the participants underwent a home interview and a nurse visit, which included the fall risk estimators used in the present study.[32] About 2 years later (wave 3), they were asked about falls experienced since the last interview.[33] Four thousand fifty-six participants aged 65 years or older concluded the interview and the nurse visit. Of those, we excluded 753 (19%) participants that at wave 3 were not reinterviewed or did not answer questions about experienced falls.

The InCHIANTI study is an observational cohort study on older adults living in the Chianti region, Italy. Its principal aim is to investigate the factors contributing to the decline of mobility in older persons and to establish clinical variables and thresholds to evaluate mobility in geriatric practice.[34] The invited persons were sampled from the municipality registries of Greve in Chianti and Bagno a Ripoli. Those aged 90 years or older were oversampled. At baseline (1999–2000), 1155 participants aged 65 years or older were assessed on a number of health parameters, including the fall risk estimators considered in the present study. After 3 years, they were re-interviewed and asked about falls experienced during the previous 12 months. We excluded 263 (23%) participants who at the first follow-up were not re-interviewed or did not answer questions about previous falls.

The Irish Longitudinal Study on Aging (TILDA) is a cohort study representative of noninstitutionalized men and women aged 50 years or older living in Ireland. It aims to study aging in Ireland in its health, economic, and social aspects.[35–37] The fieldwork relative to the baseline was carried out between October 2009 and February 2011. At baseline, the participants were asked about falls experienced during the last year and were assessed on a number of fall risk estimators. The first follow-up was carried out after about 2 years (from April 2012 to January 2013). At the follow-up, the participants were asked about falls experienced since the baseline interview. Two thousand three hundred seventy-two participants aged 65 years or older concluded the interview and the health assessment. We excluded 271 (11%) participants who at the first follow-up were not re-interviewed or did not answer questions about experienced falls. TILDA and ELSA are considered sister surveys, as both were designed similarly according to the United States Health and Retirement Study.[38]

Each of these 4 studies has received ethical approval by local competent ethics committees.

Variable Harmonization

We had to develop 4 harmonization blocks because the 4 cohort studies are different in the way they were designed and carried out. The process of deriving common variables from different existing datasets is often called “retrospective harmonization.” It allows the utilization of data coming from different sources within 1 combined analysis.[39] We call “target variables” the variables that are desired as a result of the harmonization process. We distinguish between “predictor target variables” and an “outcome target variable.” Predictor target variables are all the fall risk factors obtained as output of FRAT-up harmonization block and taken as input by the FRAT-up core block. The outcome target variable is the

object of prediction (ie, occurrence of any fall during 1 year after the assessment, hereinafter "subsequent falls"). We call "source variables" the variables that are native of each dataset and that are used to construct the target variables. Predictor source variables are the risk estimators received in input by the FRAT-up harmonization block.

For each dataset, harmonization rules were developed and applied whenever possible to construct the target variables from the source variables. This process was fully blinded, meaning that the effect of the different choices of the harmonization process on the performance of any predictive model was not evaluated.

It was considered impossible to construct 5 and 3 risk factors in the ELSA and TILDA datasets, respectively. The outcome variable was harmonized imperfectly in all the datasets except ActiFE. In the InCHIANTI dataset, the corresponding source variable is relative to a time span that comes 2 years after the assessment, whereas in the ELSA and TILDA datasets, the corresponding source variables are relative to a time span that covers 2 years instead of 1. A more detailed description of the source and target variables and of the harmonization process is provided in an Appendix that is available upon request from the corresponding author.

Statistical Analysis

Use of sample weights—In health surveys, it is often the case that the study sample, which is available for the analyses, is not fully representative for the target population. This happens because some population strata are purposely oversampled or because there can be differential response and drop-out rates. As a consequence, it may happen that the distribution of some quantities of interest in the sample population differs substantially from the distribution in the target population. Sample weights are, thus, used to make sample estimates closer to their respective target population quantities.[40]

The ELSA and TILDA datasets are released with a set of sample weights. Among those, for ELSA, we have considered the weights assigned to the participants who underwent the nurse visit. For TILDA, we have considered the weights assigned to the participants who completed the health assessment, either at home or at the health center. The weights for the samples of the ActiFE and InCHIANTI datasets were calculated after stratifying by age group and sex (for the InCHIANTI we also stratified by site, Greve in Chianti or Bagno a Ripoli[34]). More in particular, each participant in stratum h was assigned a weight N_h/n_h, with N_h (n_h) being the total number of participants in stratum h in the target population (in the available sample, respectively).

Data imputation

Missing data are less of an issue for FRAT-up because of the ability of the tool to handle missing information through use of prevalence proportions.[22] Conversely, missing data imputation is a necessary preprocessing step before computation of the data-driven models. Missing data have been imputed in 11 copies with multivariate imputations by chained equations.[41] Percentage of missing values, when different from zero, is indicated in square

brackets in Table 1. Totally missing variables in a dataset (eg, number of medications in the ELSA dataset) were replaced with the prevalence rates used in FRAT-up (Table 1).

Descriptive statistics

Descriptive statistics of the 4 cohorts were calculated for the harmonized variables using sample weights. Univariate associations between single risk factors and subsequent falls were quantified with odds ratios (ORs) and corresponding 95% confidence intervals (CIs).

Development of cohort-specific risk models

FRAT-up was applied on the 4 harmonized datasets. Its performance on the datasets was then compared with the performances of data-driven, cohort-specific risk models, estimated by 10-fold cross-validation.

In particular, each harmonized dataset was once used as a training set and, to this aim, randomly divided in 10 folds, balanced with respect to number of fallers. In turn, one of the imputed copies of 9 folds was used to fit a stepwise logistic regression with Akaike information criterion as model selection metrics.[42] All FRAT-up risk factors were included as candidate regressors, together with their 2-way interactions. This regression model was then used to calculate the risk score on the test fold of the same dataset. This procedure was repeated 10 times, to calculate risk scores on all the samples of the dataset. One randomly chosen model among these 10 was used to obtain risk scores also on the other 3 harmonized datasets used as testing sets.

To calculate risk scores, each regression model was applied on each imputed copy of the samples, obtaining 11 risk scores for each participant. These 11 scores were then averaged to obtain a unique risk score for each participant.

Model evaluation

The area under the receiver operating characteristic curve (AUC) was chosen to evaluate FRAT-up and the other cohort-specific risk models because this is the most common statistics to evaluate the discriminative ability of prediction models. Mean and 95% CIs for model AUCs were derived by means of bootstrapping.[43,44] Observations were sampled with replacement with probability proportional to their sample weights.

FRAT-up was also graphically evaluated for calibration. To draw the calibration plot, the FRAT-up 1-year risk of falling ($p1$) was adjusted to a 2-year risk of falling ($p2$) for the ELSA and the TILDA dataset according to the formula $p2 = p1(2 - p1)$. The method is further explained in the Appendix, available upon request from the corresponding author.

The values of FRAT-up AUCs attained on the 4 populations were pooled with random effects meta-analysis using the R package “meta.”[45] In particular, mean AUC was estimated with inverse variance weighted average. Between-study heterogeneity was quantified with Higgins-Thompson I^2,[46] and between-study variance with the DerSimonian-Laird estimate.[47]

Results

Table 1 describes the 4 cohorts with respect to main sociodemographic and medical characteristics as obtained after the harmonization process. Most characteristics showed a large variation and difference among the 4 cohorts, except sex, history of diabetes, and use of antiepileptics.

Table 2 reports for each cohort univariate associations of the single risk factors with risk of subsequent falls. ORs quantified in the meta-analysis by Deandrea et al[24] and used by FRAT-up are reported for comparison. As expected, most ORs are statistically significant in ELSA, InCHIANTI, and TILDA. Surprisingly, only 6 ORs are statistically significant in the ActiFE dataset. History of falls is among the strongest risk factors. In ELSA, the exceptionally high OR may be explained by the particular predictor and outcome source variables employed.

Table 3 reports the AUCs attained by FRAT-up and by the cohort-specific risk models on the 4 cohorts. The AUC of FRAT-up is 0.562 (95% CI 0.530–0.594), 0.699 (95% CI 0.680–0.718), 0.636 (95% CI 0.594–0.681), and 0.685 (95% CI 0.660–0.709) respectively, for ActiFE, ELSA, InCHIANTI, and TILDA. In each cohort, FRAT-up discriminant ability is surpassed, at most, by the cohort-specific risk model fitted on that same cohort. On the InCHIANTI cohort, FRAT-up has higher discriminative accuracy than the InCHIANTI-specific risk model.

The mean FRAT-up AUC estimated by pooling results obtained on the 4 cohorts with random effects meta-analysis is 0.646 (95% CI 0.584–0.708). The between-cohort variance is 0.0038 and the Higgins-Thompson I^2 measure of heterogeneity is 95.1% (95% CI 90.3% −97.5%). Cochran's Q-test for heterogeneity is highly significant ($P < .0001$) indicating substantial heterogeneity among the included studies (Figure 1).

Figure 2 shows the calibration curves of FRAT-up for the 4 datasets. Participants of ActiFE with low (high) risk scores, experienced more (respectively, less) falls than expected. This pattern, sometimes referred to as low resolution,[48] is also present in the participants of InCHIANTI who were assigned to the lowest or highest risk score deciles. In ELSA and TILDA, FRAT-up overestimated the risk consistently across the risk strata.

Discussion

In this comparative study, we investigated the performance of FRAT-up as a prediction tool for falls in 4 cohorts of European community-dwelling older adults, and we compared its discriminative ability with those of cohort-specific, data-driven risk models. Overall, FRAT-up seems suitable to be applied across different cohorts, thereby being a valid approach to estimate risk of falls in populations of community-dwelling older adults, although the performance varied among the different cohorts.

FRAT-up mean AUC for any fall was estimated to be 0.646 by meta-analysis of the AUCs obtained from the 4 cohorts. Compared with prediction tools for other health outcomes, such as prediction tools for cardiovascular health,[1] this value per se cannot be considered high.

However, previous research has already shown that the FRAT-up discriminative ability is superior to other screening tools,[49] such as gait speed and the Short Physical Performance Battery (SPPB).[50] Also, the Timed Up and Go test has been shown to have discriminative ability for falls similar to gait speed.[51] Its ability to predict falls[6,15] has been quantified with an AUC ranging from 0.61[52] to 0.71 (value obtained when discriminating recurrent fallers).[53] The AUC of the Tinetti Balance test[54] has been reported to be around 0.56[52] and 0.62.[55] Thus, the risk models for falls that have been proposed and validated so far, have left a conspicuous part of the phenomenon unexplained. Nevertheless, these considerations and the results of our study suggest that FRAT-up is a suitable screening tool to use in populations of community-dwelling older people.

In all cohorts, FRAT-up risk score was predictive for future falls. Without considering the results obtained when fitting and testing a model on the same population (also known as internal validation[56]), we note that for any given test cohort, the FRAT-up discriminative ability was comparable to or even greater than the other cohort-specific risk models. Furthermore, ELSA was the test cohort on which the models attained the highest and ActiFE the one with the lowest AUCs, respectively.

Besides differences among the studies in terms of risk factor prevalence rates and ORs, the I^2 statistics indicated substantial heterogeneity among the 4 included studies. It is not possible to unequivocally determine to which degree this heterogeneity is attributable to true population dissimilarities (eg, differences in the distribution of the SPPB score) or to differences in the study protocols and data collection procedures (eg, methods of recording fall occurrences). This limitation is partly due to the lack of consistent data across the studies (eg, SPPB is not available in TILDA), and to the small number of datasets included (ie, 4), therefore, not allowing to conduct a meta-regression, which might have shed further light on potential reasons. Nevertheless, first a high heterogeneity in terms of risk factors ORs was already found in the meta-analysis on which FRAT-up was built.[24] Second, some variables and the resulting heterogeneity might be the result of a sometimes imperfect harmonization process. For example, estimating exposure to the risk factor “pain” requires having a consistent and specific definition of it. However, in the actual implementation of the harmonization process, we had to deal with questionnaires being different across the 4 datasets in terms of frequency (eg, assessment of frequent or occasional pain), reference time period (eg, 12 months or 2 years before the assessment), or differences in location (eg, pain in any body location or in specific areas). Therefore, some limitations are intrinsic in the 4 different datasets; others might have been mitigated by an expert consensus process.

Other considerations to explain heterogeneity in results regard the outcome target variable (ie, occurrence of at least 1 fall in the 12 months after the assessment). First, from theoretical analyses, we expect that longer follow-ups lead to higher AUCs.[57] This may explain why, excluding results from internal validation, AUC is consistently higher on ELSA and TILDA, where participants report about falls experienced during a time period of 2 years, which is twice longer than in ActiFE and InCHIANTI. Second, the differences in the approaches used to assess fall incidence could have played a role. In particular, use of prospective falls calendars (as employed in ActiFE) is expected to be more precise,[30] whereas retrospective questionnaire assessment (as used in ELSA, InCHIANTI, and TILDA) might register only

more severe fall events, that are supposedly more easily predictable from information about exposure to standard risk factors. Finally, the differences in fall incidence among the study populations provide another potential explanation for the different behavior of FRAT-up in calibration. In particular, FRAT-up was developed assuming an average 1-year prevalence of 31% for at least 1 fall.[22] This value is similar to the prevalence of 35% found in ActiFE, where FRAT-up is substantially calibrated, whereas is much higher than the prevalence of at least 1 fall of 23% and 21% found in ELSA and InCHIANTI, respectively, where FRAT-up overestimates the risk. Discrepancies of fall incidence figures among populations is indeed a debated issue in the literature.[58,59]

Externally validating a prediction model means to evaluate the performance of the model on data that were not used for its development. It is of fundamental utility as it allows evaluating the generalizability of the model outside the derivation cohort. In addition, it allows estimating its predictive ability excluding some sources of bias that may intervene in other types of validation procedures.[60,61] External validation is rarely performed, partly because it is time-consuming and costly. Also, in the domain of falls, only few prediction models for community-dwelling older adults have been externally validated, and they have shown modest predictive accuracy.[6,16] By performing a harmonization process, that is connatural with the FRAT-up 2-block architecture, we have been able to apply and evaluate this tool on 4 datasets relative to 4 cohort studies of European older people. The issues discussed above related to the harmonization process can be thought as the price to pay for avoiding a long and expensive data collection campaign. However, if FRAT-up is conceived to be applied on multiple data sources after the construction of specific harmonization blocks, our approach to validate it reflects its intended way of using it.

Conclusions

Despite extensive research, falls are still difficult to predict because of the multiplicity of risk factors involved. Applying FRAT-up on different cohorts where risk factors were assessed according to different procedures and policies resulted in a risk score that was significantly predictive for falls, although with very heterogeneous discrimination ability. Overall, FRAT-up seems more suitable to be transferred across different cohorts than data-driven fall-risk models stemming from individual cohorts, thereby being a valid option to use on populations of community-dwelling older people if no specifically validated, population-specific fall risk tools already exist for the respective population. Nevertheless, further studies should be performed to better understand the reasons for the observed heterogeneity and to refine a tool that performs homogeneously with higher accuracy measures across different populations.

Acknowledgments

The ActiFE study was supported by grants from the Ministry of Science, Research and Arts, state of Baden-Wuerttemberg, Germany, as part of the Geriatric Competence Center, Ulm University.

The data relative to ELSA were made available through the United Kingdom Data Archive. ELSA was developed by a team of researchers based at the NatCen Social Research, University College London and the Institute for Fiscal Studies. The data were collected by NatCen Social Research. The funding is provided by the National Institute of Aging in the United States, and a consortium of United Kingdom government departments coordinated

by the Office for National Statistics. The developers and funders of ELSA and the Archive do not bear any responsibility for the analyses or interpretations presented here.

The InCHIANTI study is currently supported by a grant from the National Institute on Aging (National Institutes of Health, National Institute on Aging, Bethesda, MD) and is coordinated by the Tuscany Regional Health Agency in a partnership with the Florence Health Care Agency, the local Administrators, and the primary care physicians of Greve in Chianti and Bagno a Ripoli. The study was initially managed by the National Institute on Research and Care of the Elderly (Ancona, Italy), and it was funded by Italian Health Ministry and by a National Institutes of Health contract. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

TILDA is an interinstitutional initiative led by Trinity College Dublin. TILDA data have been co-funded by the Government of Ireland through the Office of the Minister for Health and Children, by Atlantic Philanthropies, and by Irish Life; have been collected under the Statistics Act, 1993, of the Central Statistics Office. The project has been designed and implemented by the TILDA study team, Department of Health and Children. Copyright and all other intellectual property rights relating to the data are vested in TILDA. Ethical approval for each wave of data collection is granted by the Trinity College Research Ethics Committee. TILDA data is accessible for free from the following sites: Irish Social Science Data Archive at University College Dublin http://www.ucd.ie/issda/data/tilda/; Interuniversity Consortium for Political and Social Research at the University of Michigan (http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/34315).

The research leading to these results has been partially funded by the European Union Seventh Framework Program (FP7/2007–2013) under grant agreement FARSEEING No. 2889.

The ActiFE study was supported by grants from the Ministry of Science, Research and Arts, state of Baden-Wuerttemberg, Germany, as part of the Geriatric Competence Center, Ulm University.

References

1. Berger JS, Jordan CO, Lloyd-Jones D, et al. Screening for cardiovascular risk in asymptomatic patients. J Am Coll Cardiol 2010;55:1169–1177. [PubMed: 20298922]
2. Stephan BC, Kurth T, Matthews FE, et al. Dementia risk prediction in the population: Are screening models accurate? Nat Rev Neurol 2010;6:318–326. [PubMed: 20498679]
3. Noble D, Mathur R, Dent T, et al. Risk models and scores for type 2 diabetes: Systematic review. BMJ 2011;7163:1–31.
4. Meads C, Ahmed I, Riley RD. A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. Breast Cancer Res Treat 2012;132:365–377. [PubMed: 22037780]
5. Win AK, Macinnis RJ, Hopper JL, et al. Risk prediction models for colorectal cancer: A review. Cancer Epidemiol Biomarkers Prev 2012;21:398–410. [PubMed: 22169185]
6. Gates S, Smith LA, Fisher JD, et al. Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. J Rehabil Res Dev 2008;45:1105–1116. [PubMed: 19235113]
7. Leslie WD, Lix LM. Comparison between various fracture risk assessment tools. Osteoporos Int 2014;25:1–21. [PubMed: 23797847]
8. Rubin KH, Abrahamsen B, Friis-Holmberg T, et al. Comparison of different screening tools (FRAX, OST, ORAI, OSIRIS, SCORE and age alone) to identify women with increased risk of fracture. A population-based prospective study. Bone 2013;56:16–22. [PubMed: 23669650]
9. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: What, why, and how? BMJ 2009;338.
10. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York, NY: Springer; 2008 p. 1317–1320.
11. Wyatt JJ, Altman DG. Commentary: Prognostic models: Clinically useful or quickly forgotten? BMJ 1995;311:1539–1541.
12. Rubenstein LZ. Falls in older people: Epidemiology, risk factors and strategies for prevention. Age Ageing 2006;35:ii37–ii41. [PubMed: 16926202]
13. Gillespie LD, Robertson MC, Gillespie WJ, et al. Interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev 2012;12:9.

14. Close JC, Lord SR. Fall assessment in older people. BMJ 2011;343:1.
15. Schoene D, Wu SM, Mikolaizak AS, et al. Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: Systematic review and meta-analysis. J Am Geriatr Soc 2013;61:202–208. [PubMed: 23350947]
16. Barry E, Galvin R, Keogh C, et al. Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: A systematic review and meta-analysis. BMC Geriatr 2014;14:14. [PubMed: 24484314]
17. Oliver D, Daly F, Martin FC, et al. Risk factors and risk assessment tools for falls in hospital in-patients: A systematic review. Age Ageing 2004;33:122–130. [PubMed: 14960426]
18. Billington J, Fahey T, Galvin R. Diagnostic accuracy of the STRATIFY clinical prediction rule for falls: A systematic review and meta-analysis. BMC Fam Pract 2012;13:76. [PubMed: 22870921]
19. da Costa BR, Rutjes AW, Mendy A, et al. Can falls risk prediction tools correctly identify fall-prone elderly rehabilitation inpatients? A systematic review and meta-analysis. PLoS One 2012;7:e41061. [PubMed: 22815914]
20. Howcroft J, Kofman J, Lemaire ED. Review of fall risk assessment in geriatric populations using inertial sensors. J Neuroeng Rehabil 2013;10:91. [PubMed: 23927446]
21. Shany T, Wang K, Liu Y, et al. Review: Are we stumbling in our quest to find the best predictor? Over-optimism in sensor-based models for predicting falls in older adults. Healthc Technol Lett 2015;2:79–88. [PubMed: 26609411]
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. [PubMed: 25693419]
23. Cattelani L, Palumbo P, Palmerini L, et al. FRAT-up Web Application. Available at: http://ffrat.farseeingresearch.eu/ Published 2014 Accessed August 24, 2016.
24. Deandrea S, Lucenteforte E, Bravi F, et al. Risk factors for falls in communitydwelling older people: A systematic review and meta-analysis. Epidemiology 2010;21:658–668. [PubMed: 20585256]
25. Collins GS, Reitsma JB, Altman DG, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. Ann Intern Med 2015;162:55–63. [PubMed: 25560714]
26. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med 2015;162:W1–W73. [PubMed: 25560730]
27. R Core Team R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing. Available at: http://www.r-project.org; 2013 Accessed August 24, 2016.
28. Denkinger MD, Franke S, Rapp K, et al. Accelerometer-based physical activity in a large observational cohort-Study protocol and design of the activity and function of the elderly in Ulm (ActiFE Ulm) study. BMC Geriatr 2010;10:50. [PubMed: 20663209]
29. Klenk J, Kerse N, Rapp K, et al. Physical activity and different concepts of fall risk estimation in older people—Results of the ActiFE-Ulm study. PLoS One 2015;10:e0129098. [PubMed: 26058056]
30. Lamb SE, Jorstad-Stein EC, Hauer K, et al. Development of a common outcome data set for fall injury prevention trials: The Prevention of Falls Network Europe consensus. J Am Geriatr Soc 2005;53:1618–1622. [PubMed: 16137297]
31. Steptoe A, Breeze E, Banks J, et al. Cohort profile: The English Longitudinal Study of Ageing. Int J Epidemiol 2013;42:1640–1648. [PubMed: 23143611]
32. Banks J, Breeze E, Lessof C, et al. Retirement, Health and Relationships of the Older Population in England: ELSA 2004 (Wave 2). London, UK: Institute for Fiscal Studies; 2006.
33. Banks J, Breeze E, Lessof C, et al. Living in the 21st century: Older people in England ELSA 2006 (Wave 3). London, UK: Institute for Fiscal Studies; 2008.
34. Ferrucci L, Bandinelli S, Benvenuti E, et al. Subsystems contributing to the decline in ability to walk: Bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. J Am Geriatr Soc 2000;48:1618–1625. [PubMed: 11129752]
35. Kearney PM, Cronin H, O'Regan C, et al. Cohort profile: The Irish Longitudinal Study on Ageing. Int J Epidemiol 2011;40:877–884. [PubMed: 21810894]

36. Cronin H, O'Regan C, Finucane C, et al. Health and aging: Development of the Irish Longitudinal Study on Ageing health assessment. J Am Geriatr Soc 2013; 61:S269–S278. [PubMed: 23662719]
37. Whelan BJ, Savva GM. Design and methodology of the Irish Longitudinal Study on Ageing. J Am Geriatr Soc 2013;61:S265–S268. [PubMed: 23662718]
38. Sonnega A, Weir DR. The Health and Retirement Study: A public data resource for research on aging. Open Heal Data 2014;2:e7.
39. Fortier I, Doiron D, Little J, et al. Is rigorous retrospective harmonization possible? Application of the DataSHaPER approach across 53 large studies. Int J Epidemiol 2011;40:1314–1328. [PubMed: 21804097]
40. Korn EL, Graubard BI. Analysis of Health Surveys. New York: Wiley; 1999.
41. van Buuren S, Groothuis-oudshoorn K. Mice: Multivariate imputation by chained equations in R. J Stat Softw 2011;45:1–67.
42. Venables VN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York: Springer; 2002.
43. Efron B Bootstrap methods: Another look at the Jackknife. Ann Stat 1979;7: 1–26.
44. Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1986;1:54–75.
45. Schwarzer G Meta: General package for meta-analysis. Available at: http://cran.r-project.org/package=meta; 2015 Accessed August 24, 2016.
46. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–1558. [PubMed: 12111919]
47. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7:177–188. [PubMed: 3802833]
48. Hsu W, Murphy AH. The attributes diagram. A geometrical framework for assessing the quality of probability forecast. Int J Forecast 1986;2:285–293.
49. Palumbo P, Palmerini L, Bandinelli S, et al. Fall risk assessment tools for elderly living in the community: Can we do better? PLoS One 2015;10:e0146247. [PubMed: 26716861]
50. 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–M94. [PubMed: 8126356]
51. Viccaro LJ, Perera S, Studenski SA. Is timed up and go better than gait speed in predicting health, function, and falls in older adults? J Am Geriatr Soc 2011;59: 887–892. [PubMed: 21410448]
52. Lin MR, Hwang HF, Hu MH, et al. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. J Am Geriatr Soc 2004;52: 1343–1348. [PubMed: 15271124]
53. Sai AJ, Gallagher JC, Smith LM, et al. Fall predictors in the community dwelling elderly: A cross-sectional and prospective cohort study. J Musculoskelet Neuronal Interact 2010;10:142–150. [PubMed: 20516631]
54. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc 1986;34:119–126. [PubMed: 3944402]
55. Raiche M, Hebert R, Prince F, et al. Screening older adults at risk of falling with the Tinetti balance scale. Lancet 2000;356:1001–1002. [PubMed: 11041405]
56. Moons KG, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio) marker. Heart 2012;98:683–690. [PubMed: 22397945]
57. Palumbo P, Palmerini L, Chiari L. A probabilistic model to investigate the properties of prognostic tools for falls. Methods Inf Med 2015;54:189–197. [PubMed: 25377164]
58. Rapp K, Freiberger E, Todd C, et al. Fall incidence in Germany: Results of two population-based studies, and comparison of retrospective and prospective falls data collection methods. BMC Geriatr 2014;14:105. [PubMed: 25241278]
59. Majdan M, Mauritz W. Unintentional fall-related mortality in the elderly: Comparing patterns in two countries with different demographic structure. BMJ Open 2015;5:e008672.
60. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98: 691–698. [PubMed: 22397946]

61. Siontis GC, Tzoulaki I, Castaldi PJ, et al. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. J Clin Epidemiol 2015;68:25–34. [PubMed: 25441703]

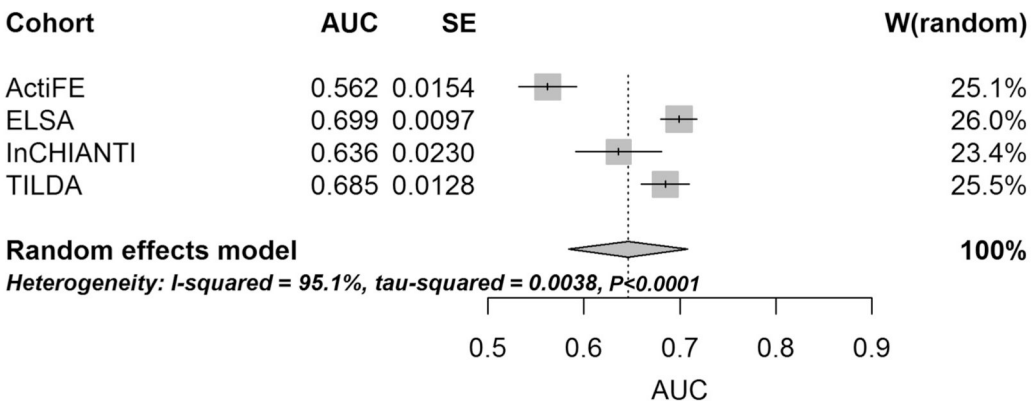

Fig. 1.

Forest plot of random effect meta-analysis for AUC attained by FRAT-up on the 4 cohorts. SE, standard error.

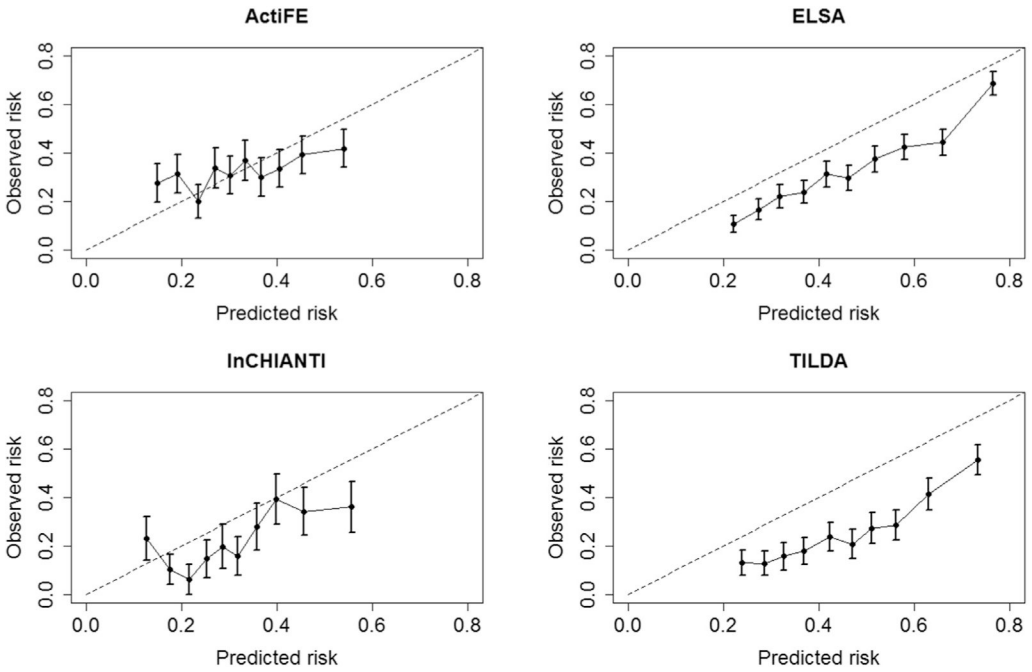

Fig. 2.

Calibration plot for FRAT-up fall risk score on the 4 datasets. The calibration curves for ELSA and TILDA are relative to falls occurred during a time span of 2 years. Error bars indicate 95% CIs.

Table 1

Main Baseline Characteristics of the 4 Included Cohorts Estimated Using Sample Weights

	ActiFE	ELSA	InCHIANTI	TILDA	Prevalence Used by FRAT-up*
Number of participants (n)	1416	3303	892	2101	
Predictor harmonized variables					
Age (years): mean (SD)	75.70 (6.76)	74.56 (7.31)	73.78 (6.62)	72.79 (5.22)	65–69 years: 25% 70–74 years: 25% 75–79 years: 20% 80–84 years: 16% 85 years+: 14%
Sex (women)	56.8%	56.7%	56.2%	53.5%	48%
1-year history of falls (yes/no)	36.1% [1.3%]	22.7% [0.4%]	20.8%	22.8% [0.05%]	31%
Living alone	27.7% [1.6%]	34.1%	18.2%	31.1%	32%
Walking aid use	1.4% [8.5%]	9.3% [1.0%]	8.2% [6.4%]	1.6%	18%
Urinary incontinence	41.0% [1.4%]	17.4% [0.1%]	34.3%	17.2% [0.5%]	19%
Diabetes mellitus	12.3% [0.3%]	10.8%	12.9%	10.5%	11%
Parkinson disease	1.6%	0.7%	1.3% [0.8%]	NA [100%]	0.8%
History of arthritis or rheumatism	52.4% [0.4%]	44.7%	34.7%	41.0%	47%
Cognition impairment (moderate to severe)	0.7% [8.1%]	0.6%	10.5%	2.5% [0.2%]	19%
History of stroke	4.9% [0.4%]	6.8%	5.8% [0.2%]	3.1%	13%
Depression (current depressive symptoms)	11.3% [5.2%]	10.0% [0.03%]	16.9% [2.5%]	3.3% [1.5%]	13%
Poor self-perceived health status	16.2% [0.5%]	NA [100%]	6.4% [2.6%]	5.9%	20%
Pain (chronic or occasional)	60.5% [0.6%]	43.1% [0.4%]	87.6% [0.4%]	39.3% [0.1%]	30%
Physical disability (difficulties in activities of daily living)	3.7% [1%]	19.3% [0.03%]	5.6%	4.6%	11%
Instrumental disability (difficulties in instrumental activities of daily living)	14.4% [1.9%]	14.4% [0.03%]	21.3%	5.1%	37%
Reported fear of falling	11.3% [1.4%]	7.5% [0.03%]	37.4% [0.1%]	32.3% [0.1%]	33%
History of dizziness	42.0% [1.0%]	22.4% [1.7%]	35.1% [6.6%]	26.5% [0.2%]	20%
Current vision impairment	83.9% [1.7%]	25.5%	54.3% [14.6%]	42.4% [23.2%]	19%
Current hearing impairment	24.2% [1.7%]	27.2%	24.1% [5.8%]	22.2%	36%
Number of medications: mean (SD)	3.62 (2.90)	NA [100%]	2.18 (2.03)	3.86 (2.87) [0.8%]	0: 23.7%, 1: 22.6%, 2: 19.4%, 3: 13.3%, 4: 8.1% 5: 4.9%, 6: 3.6%, 7: 2.0%

	ActiFE	ELSA	InCHIANTI	TILDA	Prevalence Used by FRAT-up*
					8: 1.0%, 9: 0.7%, 10: 0.7%
Use of antihypertensives	56.4%	NA [100%]	37.5%	57.7%	32%
Use of sedatives	1.3%	NA [100%]	5.7%	NA [100%]	14%
Use of antiepileptics	1.7%	NA [100%]	1.4%	NA [100%]	1%
Physical activity limitations	14.0% [14.3%]	8.1% [0.09%]	19.8% [0.3%]	36.3% [0.5%]	56%
Gait problems	22.5% [3.0%]	31.8% [8.1%]	18.4% [8.0%]	17.3% [18.4%]	42%
Outcome harmonized variable					
Subsequent falls (yes/no)	32.9%	1-year adjusted: 22.1% (2 years 33.5%)	22.8%	1-year adjusted: NA† (2 years 27.1%)	
Other characteristics					
Grip strength (kg): mean (SD)	32.18 (11.09) [1.8%]	26.38 (10.17) [2.0%]	29.92 (11.62) [18.9%]	24.00 (8.86) [0.6%]	
Gait speed (m/s): mean (SD)	0.96 (0.29) [5.3%]	0.85 (0.25) [8.5%]	1.02 (0.26) [9.9%]	NA [100%]	
SPPB balance subscore: mean (SD)	3.68 (0.81) [2.5%]	3.27 (1.24) [0.03%]	3.38 (1.13) [6.5%]	NA [100%]	
SPPB gait subscore: mean (SD)	3.59 (0.91) [3.0%]	3.47 (0.89) [8.1%]	3.67 (0.81) [8.0%]	NA [100%]	
SPPB chair standing subscore: mean (SD)	3.16 (1.16) [1.6%]	2.40 (1.45) [6.0%]	3.16 (1.22) [6.7%]	NA [100%]	
SPPB score: mean (SD)	10.45 (2.36) [5.6%]	9.46 (2.67) [13.1%]	10.23 (2.78) [8.3%]	NA [100%]	

NA, not available; SD, standard deviation.

If values were missing, percentage of missing values is indicated in square brackets.

*Values from Cattelani et al.[22]

†Not available because of lack of information about fall counts.

Table 2

Univariate Associations (ORs) With Risk of Subsequent Falls, as Estimated From the 4 Datasets and as Reported in the Meta-analysis by Deandrea et al[24]

	Odds Ratio (95% CI)				
	ActiFE	**ELSA (2-year Risk)**	**InCHIANTI**	**TILDA (2-year Risk)**	**Deandrea et al***
Age (5-year increase)	1.04 (0.95–1.13)	**1.32 (1.25–1.39)**	**1.18 (1.06–1.32)**	**1.14 (1.04–1.26)**	**1.12 (1.07–1.17)**
Sex (women)	**1.40 (1.12–1.75)**	**1.44 (1.24–1.67)**	**1.47 (1.07–2.03)**	**1.49 (1.23–1.81)**	**1.30 (1.18–1.42)**
1-year history of falls (yes/no)	**1.58 (1.25–1.99)**	**8.40 (6.68–10.55)**	**1.89 (1.33–2.69)**	**3.50 (2.82–4.34)**	**2.77 (2.37–3.25)**
Living alone	1.06 (0.82–1.38)	**1.63 (1.40–1.89)**	1.38 (0.94–2.03)	**1.46 (1.19–1.80)**	**1.33 (1.21–1.45)**
Walking aid use	1.53 (0.62–3.81)	**2.92 (2.27–3.77)**	**1.74 (1.02–2.95)**	**3.23 (1.44–7.26)**	**2.18 (1.79–2.65)**
Urinary incontinence	**1.58 (1.26–1.99)**	**1.73 (1.44–2.08)**	1.32 (0.96–1.81)	**1.61 (1.26–2.06)**	**1.40 (1.26–1.57)**
Diabetes mellitus	0.94 (0.67–1.31)	1.24 (0.99–1.56)	1.17 (0.75–1.82)	1.08 (0.79–1.48)	**1.19 (1.08–1.31)**
Parkinson disease	1.40 (0.63–3.15)	**4.48 (1.82–11.01)**	0.70 (0.15–3.26)	NA	**2.71 (1.08–6.84)**
History of arthritis or rheumatism	**1.34 (1.07–1.68)**	**1.72 (1.48–1.99)**	**1.69 (1.22–2.33)**	**1.63 (134–1.98)**	**1.47 (1.28–1.70)**
Cognition impairment (moderate to severe)	1.82 (0.47–7.13)	1.84 (0.71–4.79)	1.50 (0.94–2.39)	0.86 (0.39–1.92)	**1.36 (1.12–1.65)**
History of stroke	0.95 (0.57–1.57)	**1.90 (1.44–2.51)**	1.05 (0.51–2.18)	**2.88 (1.70–4.89)**	**1.61 (1.31–1.98)**
Depression (current depressive symptoms)	1.14 (0.79–1.65)	**1.72 (1.36–2.18)**	**2.16 (1.49–3.13)**	**1.93 (1.16–3.23)**	**1.63 (1.36–1.94)**
Poor self-perceived health status	1.23 (0.91–1.66)	NA	**2.22 (1.32–3.72)**	**2.30 (1.55–3.41)**	**1.50 (1.15–1.96)**
Pain (chronic or occasional)	1.18 (0.94–1.48)	**1.67 (1.44–1.93)**	1.51 (0.91–2.50)	**2.11 (1.73–2.56)**	**1.39 (1.19–1.62)**
Physical disability	1.52 (0.84–2.74)	**2.63 (2.19–3.15)**	**2.24 (1.23–4.08)**	**2.15 (1.36–3.40)**	**1.56 (1.22–1.99)**
Instrumental disability	1.27 (0.91–1.78)	**2.43 (1.98–2.99)**	**2.16 (1.53–3.05)**	**2.68 (1.71–4.19)**	**1.46 (1.20–1.77)**
Reported fear of falling	**1.43 (1.01–2.03)**	**3.50 (2.64–4.63)**	**1.87 (1.37–2.57)**	**2.28 (1.86–2.79)**	**1.55 (1.14–2.09)**
History of dizziness	1.12 (0.89–1.40)	**2.55 (2.15–3.03)**	1.01 (0.72–1.41)	**1.98 (1.59–2.45)**	**1.80 (1.39–2.33)**
Current vision impairment	0.95 (0.70–1.30)	**1.64 (1.39–1.94)**	**1.51 (1.05–2.17)**	1.04 (0.82–1.31)	**1.35 (1.18–1.54)**
Current hearing impairment	1.23 (0.95–1.59)	**1.37 (1.17–1.61)**	1.18 (0.83–1.67)	**1.43 (1.13–1.80)**	**1.21 (1.05–1.39)**
Number of medications (1-drug increase)	1.03 (0.99–1.07)	NA	**1.11 (1.03–1.19)**	**1.10 (1.06–1.13)**	**1.06 (1.04–1.08)**
Use of antihypertensives	1.11 (0.89–1.40)	NA	NA	1.09 (0.90–1.32)	**1.25 (1.06–1.48)**
Use of sedatives	0.52 (0.17–1.56)	NA	NA	NA	**1.38 (1.15–1.66)**
Use of antiepileptics	**2.63 (1.25–5.52)**	NA	NA	NA	**1.88 (1.02–3.49)**
Physical activity limitations	1.11 (0.79–1.56)	**2.14 (1.63–2.79)**	**2.44 (1.70–3.49)**	**1.36 (1.11–1.66)**	**1.20 (1.04–1.38)**
Gait problems	1.02 (0.77–1.35)	**1.94 (1.64–2.28)**	**2.39 (1.65–3.47)**	**1.55 (1.13–2.11)**	**2.06 (1.82–2.33)**

Statistical significant ORSs are reported in bold.

*Values from Deandrea et al.[24]

Table 3

Comparison Among Models Applied on the 4 Cohorts

	AUC (95% CI)			
	ActiFE	**ELSA**	**InCHIANTI**	**TILDA**
FRAT-up	0.562 (0.530–0.594)	0.699 (0.680–0.718)	0.636 (0.594–0.681)	0.685 (0.660–0.709)
Cohort-specific model fitted on ActiFE	0.574 (0.541–0.604)	0.566 (0.545–0.585)	0.549 (0.505–0.594)	0.559 (0.532–0.584)
Cohort-specific model fitted on ELSA	0.560 (0.527–0.593)	0.719 (0.698–0.739)	0.611 (0.570–0.654)	0.675 (0.648–0.704)
Cohort-specific model fitted on InCHIANTI	0.530 (0.501–0.559)	0.664 (0.644–0.681)	0.571 (0.520– 0.619)	0.633 (0.608–0.661)
Cohort-specific model fitted on TILDA	0.561 (0.527–0.592)	0.661 (0.642–0.678)	0.600 (0.558–0.647)	0.686 (0.660–0.710)

The discriminative ability is quantified with AUC (95% CI). The results from internal validation (fitting and testing on the same cohort) are in italics.