Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

BIOINGEGNERIA

Ciclo XXVII

Settore Concorsuale di afferenza: 09/G2
Settore Scientifico disciplinare: ING-INF/06

BIOMEDICAL ENGINEERING FOR HEALTHY AGEING. PREDICTIVE TOOLS FOR FALLS

Presentata da: **Pierpaolo Palumbo**

Coordinatore Dottorato **Relatore**

Prof.ssa Elisa Magosso **Prof. Lorenzo Chiari**

Esame finale anno 2015

Abstract

Falls are common and burdensome accidents among the elderly. About one third of the population aged 65 years or more experience at least one fall each year. Fall risk assessment is believed to be beneficial for fall prevention. This thesis is about prognostic tools for falls for community-dwelling older adults.

We provide an overview of the state of the art. We then take different approaches: we propose a theoretical probabilistic model to investigate some properties of prognostic tools for falls; we present a tool whose parameters were derived from data of the literature; we train and test a data-driven prognostic tool. Finally, we present some preliminary results on prediction of falls through features extracted from wearable inertial sensors.

Heterogeneity in validation results are expected from theoretical considerations and are observed from empirical data. Differences in studies design hinder comparability and collaborative research. According to the multifactorial etiology of falls, assessment on multiple risk factors is needed in order to achieve good predictive accuracy.

Chapter 1

Introduction

FALL PREDICTION

Chapter 2

Theoretical model

Chapter 3

Learning from the literature
Tests across datasets

Chapter 4

Learning from data

Chapter 5

Prediction from wearable inertial sensors

Chapter 6

Conclusion

Table of contents

Chapter 1

This first chapter introduces the theme of prediction in medicine in general, and more in particular it gives an overview of fall risk screening and prediction algorithms suggested by clinical guidelines and presented in the literature. It closes with a presentation of the following chapters.

Ageing

Declining mortality and fertility rates are shaping the demographic structure of both developed and less developed regions all around the world. It is estimated that during the last 200 years life expectancy at birth is doubled, growing at a pace of two years per decade [1]. Since 1950 total fertility rate has halved, from 5 children per woman to the current rate of 2.5 [2]. Population ageing has major consequences of economic, social and political nature. From an epidemiological perspective, population ageing is contributing to the shift of disease burden from communicable, neonatal and maternal diseases to non-communicable diseases and injuries [3].

Falls

Falls are common and burdensome accidents among the older population. About one third of the population aged 65 or more experience at least one fall each year [4] and the annual fall rate in this age group is about 0.65 falls per person [5]. Fall risk increases with age and is higher in populations of older adults hospitalized or living in long-term care institutions. Falls can result in injuries and are a leading cause of activity restriction, hospitalization, and disability [6; 7]. Physical injuries due to falls account for 40% of all injury deaths [4]. Worldwide, it is estimated that falls are responsible for 35 million disability adjusted life years [3].

Many preventive strategies have been proposed, and some of them have been shown to be effective [8–11]. Their implementation, however, has been slow and the coverage insufficient [12–14]. The individual and societal costs of these interventions are often among the factors that hinder their implementation. In order to make best use of available resources and intervene only with subjects at increased risk, medical associations and national health authorities recommend the adoption of fall risk assessment tools [15–19].

The expression “fall risk assessment tool” is used in the literature with two different acceptations. Indeed, sometimes fall risk assessment tool refers to screening tests whose aim is only to identify subjects at increased risk. Other times, by “fall risk assessment” it is meant the act of thoroughly assessing the presence of risk factors to target with interventions. The most widely accepted paradigm for fall prevention

in community-dwelling older adults encompasses both. In particular, it consists of three sequential stages: screening for high fall risk; assessment of those at high risk on multiple risk factors; implementation of a tailored intervention [11; 15; 18; 20].

Prediction in medicine

Prognostic tools

Let us consider the general case of a user that wants to have information about a health outcome of a particular subject (the subject under assessment). By *prediction tool* we mean any tool that receives information about the subject under assessment –and possibly other contextual information, e.g. environmental information– and gives the user information about the health outcome of interest. When the information that is provided by the tool is about the occurrence of the outcome in the future, we also use the term '*prognostic tool*' (or *prognostic model*).

Their objectivity, i.e. their characteristic of providing the output in a user-independent manner, given the input information, is of value for evidence-based medicine.

Their employment is diffuse in clinical practice, public health, and medical research [21–23]. In clinical practice, they inform the patient and the physician about the possibility of the occurrence of an adverse event (e.g. death or onset of a disease) or of the success or failure of a given therapy; they support decisions about the appropriateness of taking further diagnostic tests or beginning a treatment; they help communication between physician and patient, and among physicians. In public health, they enable policies that target preventive interventions only to subjects at increased risk; they allow comparison between and assessment of healthcare providers (e.g. hospitals) adjusting for different case-mix. In medical experimental research, they are employed in the design (for patient recruitment [24] and stratification) and analysis (adjustment for random imbalance, increase in statistical power) of randomized controlled trials. In observational studies investigating the effect of a given treatment, they are used as propensity scores (scores expressing the probability of receiving the treatment) to adjust for the potential confounding factor of differential treatment.

Overview of results in the literature

Though the focus of this thesis is on predictive tools for falls, we like to give a broad and synthetic overview of the results that have been published in the literature about prediction of different health outcomes.

Figure 1 shows the discriminative ability of different prognostic models for the onset of different non-communicable diseases and the occurrence of accidents. The discriminative ability of the different models has been quantified with the Area Under the ROC Curve (AUC) or with different definitions of C-statistics for survival models in the presence of censoring [25; 26]. We included only results obtained from external validation studies. More details about methods employed to obtain this figure, references and more information about these models are included in Appendix 1.

When drawing comparisons, it should be borne in mind that these discrimination indices were obtained on different populations and from follow-ups of different durations. As an example, the follow-ups in studies about cardiovascular events are generally of 5-10 years, whereas a common follow-up duration for studies about falls is 6-12 months. We can define a prediction task as the task of predicting the occurrence of a particular kind of event (e.g. cardiovascular event) over an associated time-span (e.g. 10 years).

The discriminative-ability values vary across the models, the studies, and even the subpopulations considered in a single study. They span from 0.5 to 0.84. However, there seems to be an effect that is

related to the object of prediction. The AUCs of models predicting cardiovascular events range between 0.7 and 0.8, whereas the AUCs obtained on models for falls seldom and barely surpass 0.6. The figure thus suggests that the actual medical knowledge allows fulfilling some prediction tasks better than others, with falls being among the most difficult objects of prediction to deal with.

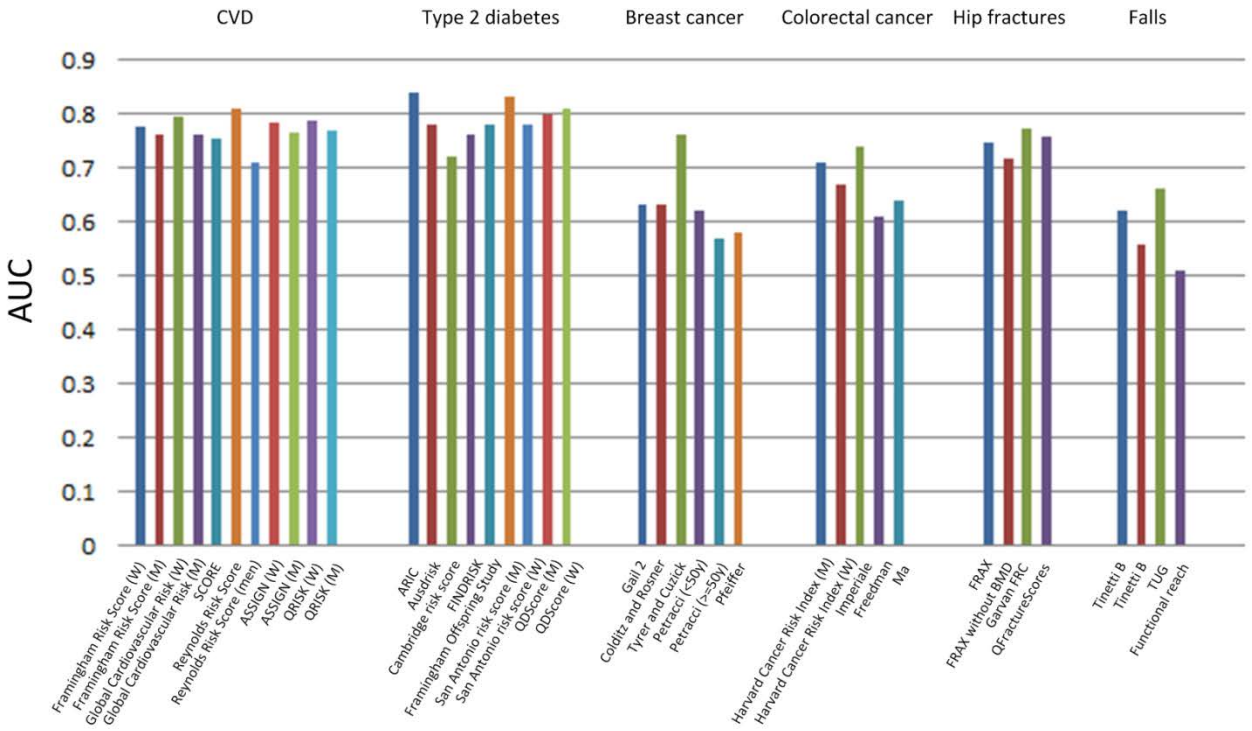

Figure 1. Area Under the Curve (AUC) or c-statistic of different prognostic models for the onset of different non-communicable diseases and the occurrence of accidents. More details in Appendix 1. CVD: cardiovascular disease.

New biomarkers

As the technology has progressed, it has offered the possibility of measuring new quantities, often quickly and cheaply. Consequently, the space of candidate predictors for given health outcomes has progressively enlarged.

This 'high-throughput' revolution and the rush to the discovery of new predictors have in turn challenged statistics with problems related to research in high-dimensional spaces (when the number of features is high and often much higher than the number of statistical units) [27].

New biomarkers have been found among DNA sequence variations, differential expression of genes, metabolites within the bloodstream, complex structures in bio-signals and bio-images, etc. The improvements over predictive tools based on traditional clinical variables have often been judged marginal, especially for common, complex health conditions [28; 29]. Since this qualitative judgment depends on the metrics used to evaluate this improvement and on personal expectations, research in biostatistics has worked to find new ways to measure the added value of new markers. With this regard, incremental AUC has been deemed not to be sensible enough for this purpose, and new metrics based on reclassification of subjects among risk strata have been introduced [30–32].

Also research on falls witnesses a 'high-throughput' revolution, represented by the employment of wearable inertial sensors. Whether general consensus has been reached on some socio-demographic, clinical and physiological risk factors, expressed in the form of systematic reviews [33–40], research on new

markers, especially features of movement analysis coming from wearable inertial sensors, is still in its early stages.

Fall risk assessment

In this paragraph we review the major tools that have been proposed for fall risk assessment. In particular, we list some algorithms issued within guidelines by national and international health authorities, and other predictive models proposed in the literature.

Guidelines

The American Geriatric Society and the British Geriatric Society issued the last update of the guideline for fall prevention in the elderly in 2011 [15]. This guideline is intended to assist healthcare professionals when visiting community-dwelling older adults in clinical setting. Figure 2 presents the proposed algorithm. It encompasses assessment and intervention. Get Up and Go test, Timed Up and Go Test (TUG), Berg Balance test, and the Performance-Oriented Mobility Assessment (also known as Tinetti Scale) [41–44] are proposed for the evaluation of gait and balance. The evidence about strengths and limitations of these tests are annotated. A thorough multifactorial fall risk assessment is suggested for subjects having abnormalities in gait or balance, having experienced two or more falls in the last year, or presenting with an acute fall.

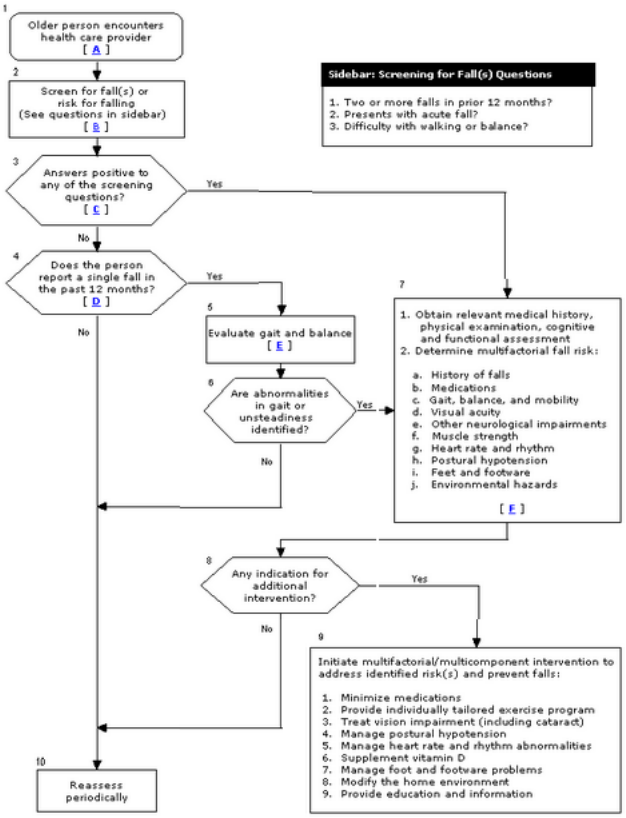

Figure 2. Algorithm from the guidelines of the American Geriatric Society and the British Geriatric Society [15; 45]**.**

The American Centers for Disease Control and Prevention (CDC) published the STEADI (Stopping Elderly Accidents, Deaths & Injuries) Tool Kit for Health Care Providers [46; 47]. This encompasses brochures for older adults and for healthcare professionals. Among those, there is an algorithm for fall risk assessment,

reported in Figure 3[1], and fall risk checklists for patients and physicians. TUG, 30 Second Chair Stands, and 4 stage balance test are recommended or suggested for evaluating gait, balance and muscle strength.

Furthermore, the CDC and the Department of Health & Human Services also published a checklist for home environmental hazards [48].

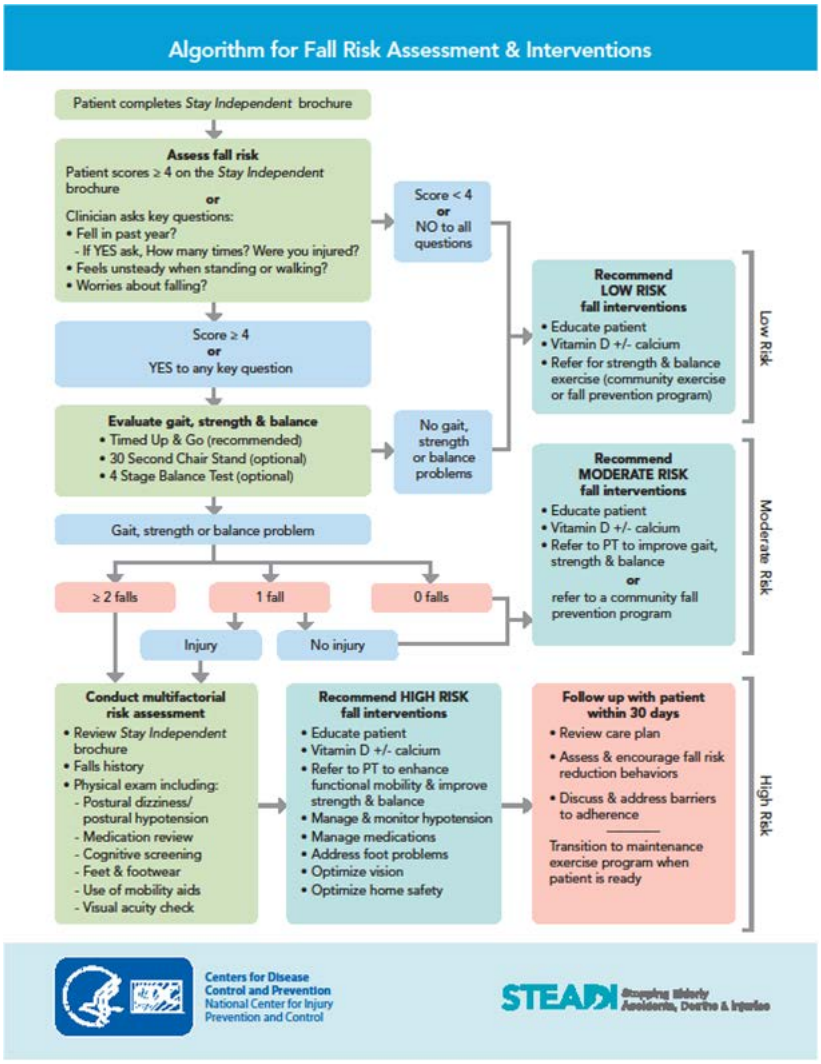

Figure 3. Algorithm for fall risk assessment and intervention issued by the CDC [49].

The English National Institute for Health and Care Excellence (NICE) published in June 2013 the guideline CG161, extending CG21 published in 2004 [50]. Figure 4 presents a flow chart thereof. The guideline explicitly advices against the use of predictive tools for falls to evaluate inpatients' risk of falling in the hospital. There is the advice to assess community-dwelling older adults reporting a fall or considered at risk of falling for their walk and balance abilities. It is left to the healthcare professional to judge when to consider a subject at risk of falling, and it is not clear whether it is specified or not how to assess gait and balance.

[1] We note that in the Algorithm for Fall Risk Assessment and Intervention proposed by the CDC there is a mistake. The "or" in "Score<4 or NO to all questions" should be replaced by "and" in order to exclude the possibility to assign one subject to two different risk strata

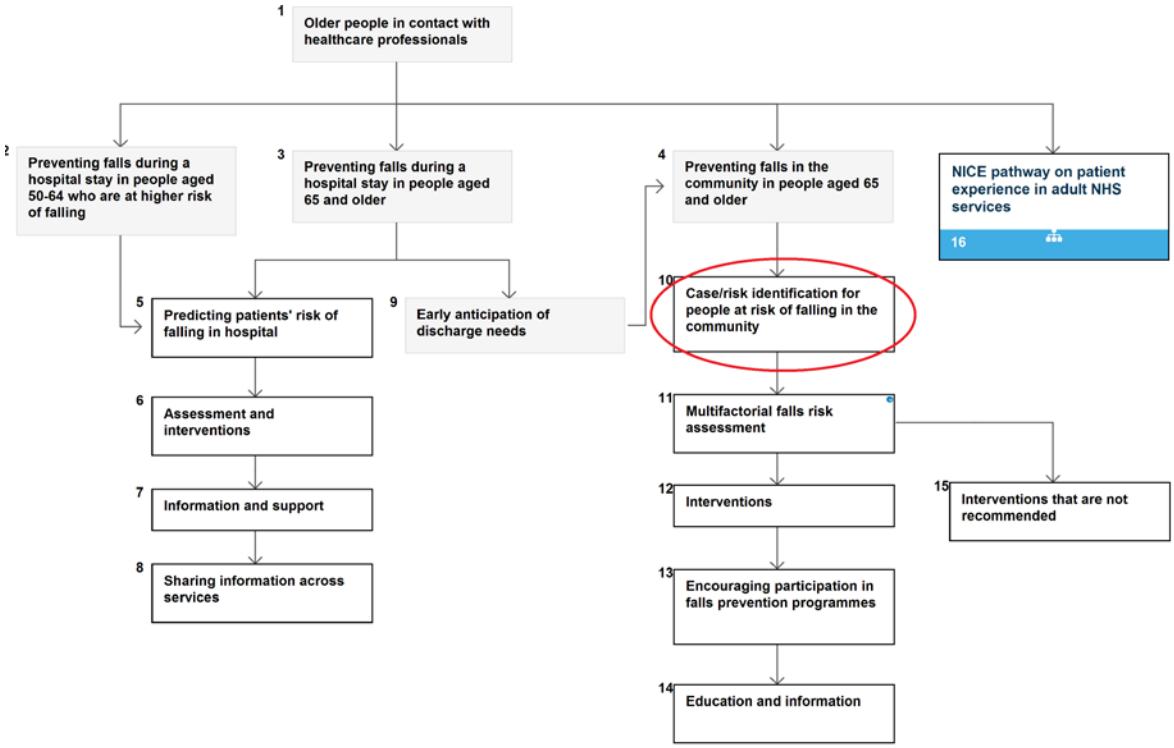

Figure 4. Flow chart issued by the NICE. Modified from [51]**.**

Following the scheme of the NICE guideline, the Italian Istituto Superiore di Sanità issued in 2007 and updated in 2009 a guideline for the prevention of falls at home for the elderly. In order to evaluate the risk of falling, the guideline recommends tests that are reliable, easy and quick to administer. The Berg Balance Scale, the TUG, and the Tinetti Balance are identified as having these characteristics. These indications are labelled with maximum degree of strength of recommendation and maximum degree of evidence [16].

More guidelines are collected in the website of ProFouND (Prevention of Falls Network for Dissemination) [52].

Literature

In the literature there are several reviews about tools for fall risk assessment for community-dwelling older adults [53; 54], for older inpatients [55–57], or in general [58–62]. Other reviews are about specific tools, e.g. the TUG [63; 64] and, the STRATIFY [65; 66]. Three reviews are about tools employing features from inertial sensors [67–69].

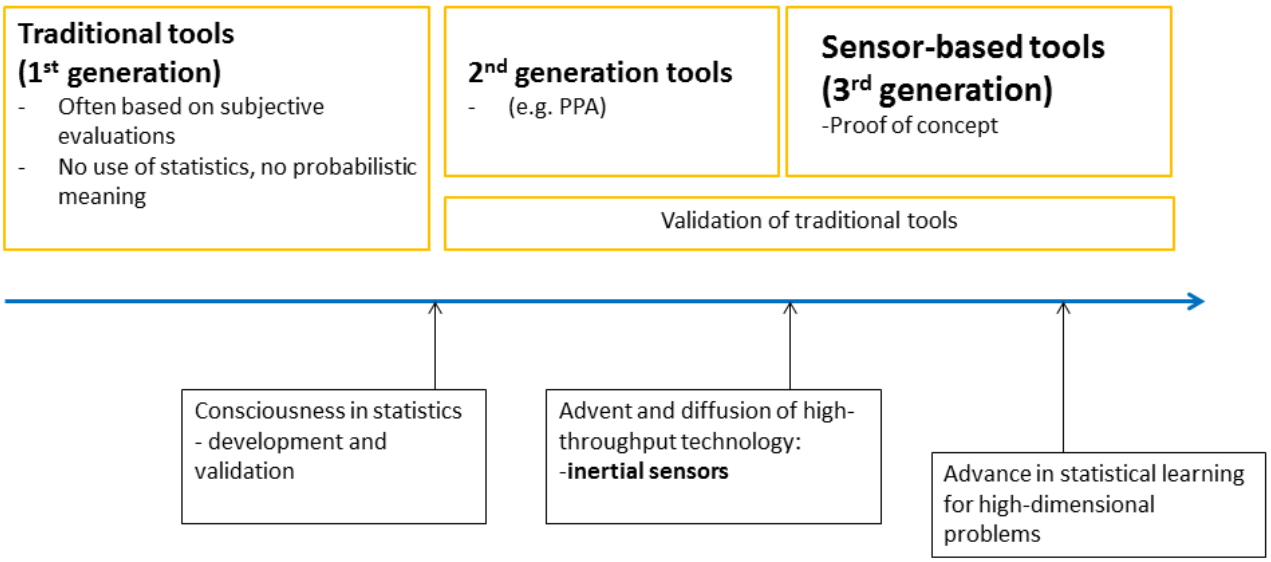

Figure 5. Schematic representation of the literature about tools for fall risk assessment. The blue arrow is the timeline. Text boxes below the timeline highlight progress in technology (inertial sensors) and information technology (statistical learning, biostatistics) that has influenced research on fall risk assessment.

Despites all its limitations, Figure 5 may be useful to have a quick and synthetic look at the literature. The first tools were published in the late '80s and during the '90s. Sometimes they were based on subjective evaluations, were developed for a more general scope (assessment of functional mobility) and without use of statistics, and gave as output a qualitative score that had no probabilistic meaning. We call these first tools 'traditional'. They are so far the ones that have been more extensively validated. Some results from external validation studies about these traditional tools are summarized in Figure 6 in terms of points and lines on the Receiver Operating Characteristic (ROC) plane. Results are from [70–74].

After many years of validation, the results about these traditional tools have been found not to be satisfactory. The TUG was proposed as a test for functional mobility in 1991 [42], modifying the Get-Up and Go test by Mathias et al [41]. In 2014 Barry et al concluded: "TUG should no longer be used as a falls risk assessment in community dwelling elderly people" [63]. Similarly, the STRATIFY (St Thomas's risk assessment tool in falling elderly inpatients) was proposed in 1997 as a tool for inpatients [75]. In 2012 Billington et al concluded: "the diagnostic accuracy of the STRATIFY rule is limited and should not be used in isolation for identifying individuals at high risk of falls in clinical practice" [66].

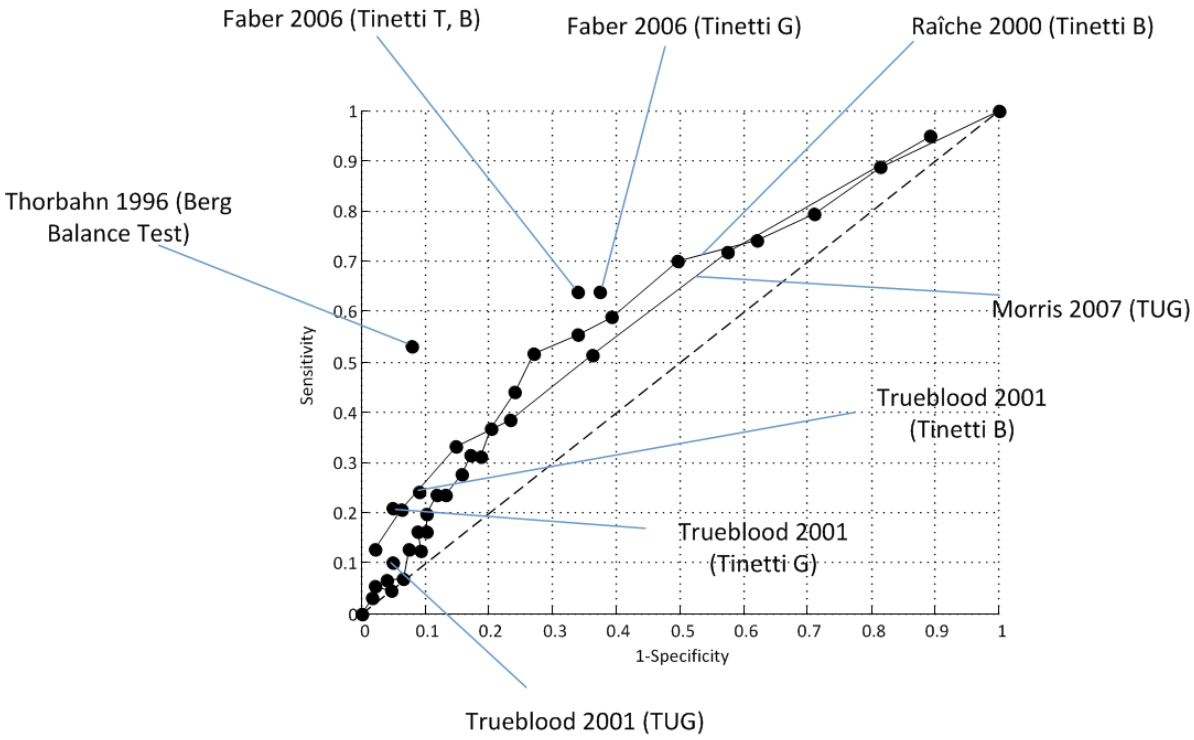

Figure 6. Sensitivity and specificity of TUG, Tinetti Balance (B), Tinetti Gait (G), Tinetti total score (T), and Berg Balance Test.

During the last decades, the adoption of results achieved in machine learning and biostatistics has improved the practice about how to develop and validate predictive tools. In the meanwhile, research on falls has gained more attention. Accordingly, after the first traditional tools, others have been published, making use of available methodological improvements (e.g. [76–78]). Among these tools –that we call 'second generation tools'– it is worth mentioning the well-known Physiological Profile Assessment (PPA) [79]. This tools has been extensively used, as a support to design and analyze clinical trials and for observational studies (e.g. [80; 81]). However, to the best of our knowledge, its ability to predict falls has not been assessed thoroughly via external validation.

The present period is dominated by research on 'sensor-based' tools –that we also call 'third generation tools'– i.e. tools that assess the risk of falling, receiving in input features extracted from signals recorded by wearable inertial sensors. These sensors have been proposed to increase the accuracy in predicting future falls while guaranteeing the objectivity of the assessment. Figure 7 presents the results on the ROC planes of these sensor-based tools, with a comparison with what achieved by validated traditional tools. Details about the studies included in this figure are given in Appendix 1.

Some issues about this research area have been already introduced above (New biomarkers). Here we add that most of these studies have been based on falls assessed retrospectively, and on small sample sizes. Furthermore, to the best of our knowledge, none of the proposed tools has been externally validated yet. Thus, the results achieved so far cannot be considered definitive or robust.

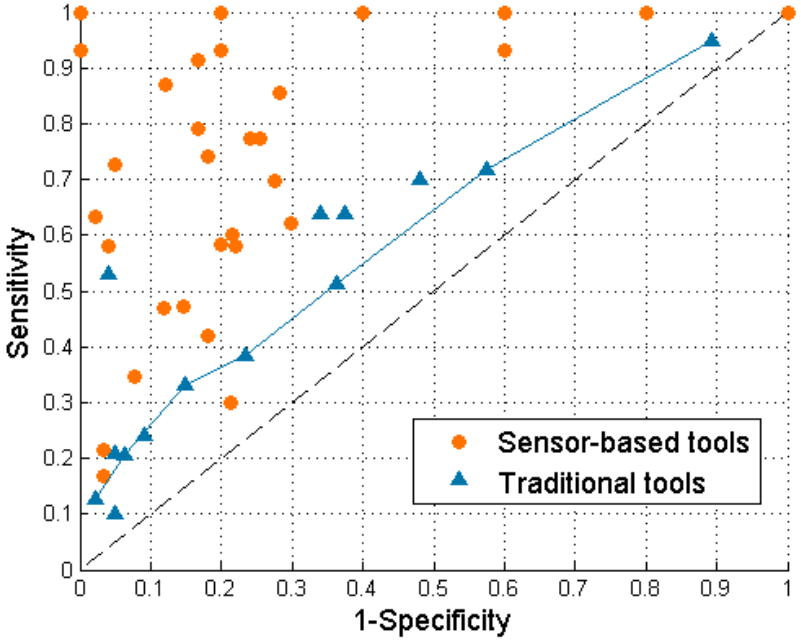

Figure 7. Sensitivity and specificity of sensor-based tools (orange circles) and externally validated traditional tools (blue triangles). More details are given in Appendix 1.

Impact

Prognostic models are developed to improve efficiency and safety of care. Impact studies are intended to evaluate this effect. Their most preferable study design is a randomized control trial, even though other designs are possible. Despite their informative value and the repeated pleas for them, their paucity is a hallmark of medical prognostic research [82–85].

To the best of our knowledge no impact study has ever been carried out for prognostic tools for falls in community-dwelling older adults. Two factors may have hindered these kinds of studies: the absence of a valid and accurate tool for fall risk, and the cost of conducting clinical trials.

Meyer et al. led a cluster-randomized control trial assessing the effectiveness of a fall risk assessment tool, the Downton Index, in nursing homes [86; 87]. They found no significance difference between the intervention and the control group. However, some points are worth discussing. The primary endpoint was the number of participants experiencing at least one fall during a 12-month follow-up. Prevention was recommended through an educational event delivered in the same way to nurse staff of both groups. Nurse staff of the intervention group was given no specific instructions on how to take advantage of the outcome of the Downton Index within the preventive initiative. Thus, it seems that in the control group a treat-everyone strategy was implemented, whereas in the intervention group the same strategy was accompanied by the use of the Downton Index. With this regard, their findings are hardly unexpected. Reilly and Evans define and suggest two measures –safety and efficiency– to consider as endpoints of impact studies [83].

Despite the absence of clinical trials dedicated to the study of the impact of prognostic tools for falls in community dwellings, some subgroup analyses from randomized control trials are encouraging. Within a Cochrane review, Gillespie et al found that home modification interventions are particularly effective in subjects at increased risk, and, in general, targeting specific high-risk groups increases the cost-effectiveness of preventive programs [9]. Similar results were found also in other studies [88].

Some difficulties of rigorous impact studies may be overcome with a modelling approach (for an example in another research area see e.g. [89]). Decision analytic models have been developed to evaluate the

economic impact of fall preventive interventions [90–93]. However, to the best of our knowledge, they did not estimate the added value of using prognostic tools for falls.

Introduction to the following chapters

The following chapters will present original contributions about predictive tools for falls.

Chapter 2 will clarify some conceptual issues that arise when assessing the performance of a probabilistic tool for fall prediction. We will employ as a model an ideal, perfect probabilistic tools that issues its forecasts on a population where risk and falls are distributed according to the classical Greenwood and Yule model for accident proneness [94]. We will show that a perfect probabilistic tool does not reach perfect discrimination, and that its performance indices are sensible to a number of key factors.

Chapter 3 will present the development of a predictive tool for falls from the literature, FRAT-up, and its validation on three populations.

In Chapter 4 will present an extensive search over a large dataset (InCHIANTI study, 2313 samples relative to 976 subjects, 1010 variables), training and validating a model that issues probabilistic predictions on the number of future falls. The model is evaluated on prospective falls and it is benchmarked against other fall risk scores: history of falls, gait speed, Short Physical Performance Battery, and FRAT-up. We study the tradeoff between two competing requirements for a prognostic tool: to be accurate in its predictions, and to be easy and quick to administer.

Chapter 5 will present some results about prediction of falls via wearable sensors obtained on data from InCHIANTI-FARSEEING [95].

In Chapter 6 we will give some final remarks.

References

1. Oeppen J, Vaupel JW. Broken limits to life expectancy. Science. 2002;296 (May):1029–1031.

2. United Nations, Department of Economic and Social Affairs, Population Division. World population ageing 2013. 2013. http://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2013.pdf. Accessed 2015-2-28.

3. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Bahalim AN, Barker-Collo S, Barrero LH, Bartels DH, Basáñez M-G, Baxter A, Bell ML, Benjamin EJ, Bennett D, Bernabé E, Bhalla K, Bhandari B, Bikbov B, Bin Abdulhak A, Birbeck G, Black JA, Blencowe H, Blore JD, Blyth F, Bolliger I, Bonaventure A, Boufous S, Bourne R, Boussinesq M, Braithwaite T, Brayne C, Bridgett L, Brooker S, Brooks P, Brugha TS, Bryan-Hancock C, Bucello C, Buchbinder R, Buckle G, Budke CM, Burch M, Burney P, Burstein R, Calabria B, Campbell B, Canter CE, Carabin H, Carapetis J, Carmona L, Cella C, Charlson F, Chen H, Cheng AT-A, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahiya M, Dahodwala N, Damsere-Derry J, Danaei G, Davis A, De Leo D, Degenhardt L, Dellavalle R, Delossantos A, Denenberg J, Derrett S, Des Jarlais DC, Dharmaratne SD, Dherani M, Diaz-Torne C, Dolk H, Dorsey ER, Driscoll T, Duber H, Ebel B, Edmond K, Elbaz A, Ali SE, Erskine H, Erwin PJ, Espindola P, Ewoigbokhan SE, Farzadfar F, Feigin V, Felson DT, Ferrari A, Ferri CP, Fèvre EM, Finucane MM, Flaxman S, Flood L, Foreman K, Forouzanfar MH, Fowkes FGR, Fransen M, Freeman MK, Gabbe BJ, Gabriel SE, Gakidou E, Ganatra HA, Garcia B, Gaspari F, Gillum RF,

Gmel G, Gonzalez-Medina D, Gosselin R, Grainger R, Grant B, Groeger J, Guillemin F, Gunnell D, Gupta R, Haagsma J, Hagan H, Halasa YA, Hall W, Haring D, Haro JM, Harrison JE, Havmoeller R, Hay RJ, Higashi H, Hill C, Hoen B, Hoffman H, Hotez PJ, Hoy D, Huang JJ, Ibeanusi SE, Jacobsen KH, James SL, Jarvis D, Jasrasaria R, Jayaraman S, Johns N, Jonas JB, Karthikeyan G, Kassebaum N, Kawakami N, Keren A, Khoo J-P, King CH, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Laden F, Lalloo R, Laslett LL, Lathlean T, Leasher JL, Lee YY, Leigh J, Levinson D, Lim SS, Limb E, Lin JK, Lipnick M, Lipshultz SE, Liu W, Loane M, Ohno SL, Lyons R, Mabweijano J, MacIntyre MF, Malekzadeh R, Mallinger L, Manivannan S, Marcenes W, March L, Margolis DJ, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGill N, McGrath J, Medina-Mora ME, Meltzer M, Mensah GA, Merriman TR, Meyer A-C, Miglioli V, Miller M, Miller TR, Mitchell PB, Mock C, Mocumbi AO, Moffitt TE, Mokdad AA, Monasta L, Montico M, Moradi-Lakeh M, Moran A, Morawska L, Mori R, Murdoch ME, Mwaniki MK, Naidoo K, Nair MN, Naldi L, Narayan KMV, Nelson PK, Nelson RG, Nevitt MC, Newton CR, Nolte S, Norman P, Norman R, O'Donnell M, O'Hanlon S, Olives C, Omer SB, Ortblad K, Osborne R, Ozgediz D, Page A, Pahari B, Pandian JD, Rivero AP, Patten SB, Pearce N, Padilla RP, Perez-Ruiz F, Perico N, Pesudovs K, Phillips D, Phillips MR, Pierce K, Pion S, Polanczyk G V, Polinder S, Pope CA, Popova S, Porrini E, Pourmalek F, Prince M, Pullan RL, Ramaiah KD, Ranganathan D, Razavi H, Regan M, Rehm JT, Rein DB, Remuzzi G, Richardson K, Rivara FP, Roberts T, Robinson C, De Leòn FR, Ronfani L, Room R, Rosenfeld LC, Rushton L, Sacco RL, Saha S, Sampson U, Sanchez-Riera L, Sanman E, Schwebel DC, Scott JG, Segui-Gomez M, Shahraz S, Shepard DS, Shin H, Shivakoti R, Singh D, Singh GM, Singh JA, Singleton J, Sleet DA, Sliwa K, Smith E, Smith JL, Stapelberg NJC, Steer A, Steiner T, Stolk WA, Stovner LJ, Sudfeld C, Syed S, Tamburlini G, Tavakkoli M, Taylor HR, Taylor JA, Taylor WJ, Thomas B, Thomson WM, Thurston GD, Tleyjeh IM, Tonelli M, Towbin JA, Truelsen T, Tsilimbaris MK, Ubeda C, Undurraga EA, van der Werf MJ, van Os J, Vavilala MS, Venketasubramanian N, Wang M, Wang W, Watt K, Weatherall DJ, Weinstock MA, Weintraub R, Weisskopf MG, Weissman MM, White RA, Whiteford H, Wiebe N, Wiersma ST, Wilkinson JD, Williams HC, Williams SRM, Witt E, Wolfe F, Woolf AD, Wulf S, Yeh P-H, Zaidi AKM, Zheng Z-J, Zonies D, Lopez AD, AlMazroa MA, Memish ZA. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec;380 (9859):2197–223. http://www.thelancet.com/journals/a/article/PIIS0140-6736(12)61689-4/fulltext. Accessed 2014-2-19.

4. World Health Organization. WHO global report on falls prevention in older age. Geneva, Switzerland: World Health Organization; 2007.

5. Rubenstein LZ. Falls in older people: epidemiology, risk factors and strategies for prevention. Age Ageing 2006 Sep;35 Suppl 2:ii37–ii41. http://www.ncbi.nlm.nih.gov/pubmed/16926202. Accessed 2012-7-13.

6. Gill TM, Allore HG, Holford TR, Guo Z. Hospitalization, restricted activity, and the development of disability among older persons. JAMA 2004 Nov;292 (17):2115–24. http://jama.jamanetwork.com/article.aspx?articleid=199719#REF-JOC41058-34. Accessed 2014-1-4.

7. Mathers LJ, Weiss HB. Incidence and Characteristics of Fall-related Emergency Department Visits. Acad. Emerg. Med. 1998 Nov;5 (11):1064–1070.

8. Kannus P, Sievänen H, Palvanen M, Järvinen T, Parkkari J. Prevention of falls and consequent injuries in elderly people. Lancet 2005 Nov;366 (9500):1885–93. http://www.ncbi.nlm.nih.gov/pubmed/16310556. Accessed 2014-1-4.

9. Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, Lamb SE. Interventions for preventing falls in older people living in the community. Cochrane database Syst. Rev. 2012; (9) http://www.ncbi.nlm.nih.gov/pubmed/22972103. .

10. Tinetti ME, Baker DI, McAvay G, Claus EB, Garrett P, Gottschalk M, Koch ML, Trainor K, Horwitz RI. A multifactorial intervention to reduce the risk of falling among elderly people living in the community. N.

Engl. J. Med. 1994 Sep;331 (13):821–7. http://www.ncbi.nlm.nih.gov/pubmed/8078528. Accessed 2013-7-6.

11. Tinetti ME. Preventing Falls in Elderly Persons. N. Engl. J. Med. 2003;348 (1):42–49.

12. Royal College of Physicians, Treml J, Husk J, Lowe D, Vasilakis N. Falling standards, broken promises. Report of the national audit of falls and bone health in older people 2010. London: 2010.

13. Royal College of Physicians. Report of the 2011 inpatient falls pilot audit. London: 2012.

14. Center of Research Excellence in Patient Safety. An evaluation of the preventing falls and harm from falls in older people best practice guidelines for Australian hospitals. 2012.

15. Panel of Falls in Older Persons American Geriatrics Society and British Geriatrics Society. Summary of the Updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. J. Am. Geriatr. Soc. 2011 Jan;59 (1):148–57. http://www.ncbi.nlm.nih.gov/pubmed/21226685. Accessed 2012-3-14.

16. Brigoni P, De Masi S, Di Franco M, Rinonapoli G, Romanini E, Sampaolo L, Torre M, Turchetti G, Vanacore N. Prevenzione delle cadute da incidente domestico negli anziani | SNLG-ISS. 2007. http://www.snlg-iss.it/lgn_cadute_da_incidente_domestico_negli_anziani. Accessed 2013-8-23.

17. Beauchet O, Dubost V, Revel Delhom C, Berrut G, Belmin J. How to manage recurrent falls in clinical practice: guidelines of the French Society of Geriatrics and Gerontology. J. Nutr. Health Aging 2011 Jan;15 (1):79–84. http://www.ncbi.nlm.nih.gov/pubmed/21267524. .

18. U.S. Preventive Services Task Force. Guide to Clinical Preventive Services, 2014. 2014. http://www.ahrq.gov/professionals/clinicians-providers/guidelines-recommendations/guide/section2a.html#falls. Accessed 2015-3-2.

19. Cattelani L, Palumbo P, Palmerini L, Bandinelli S, Becker C, Chesani F, Chiari L. FRAT-up, a web-based fall risk assessment tool for elderly people living in the community. J. Med. Internet Res. 2015;17 (2).

20. Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ. Will my patient fall? JAMA 2007 Jan;297 (1):77–86. http://www.ncbi.nlm.nih.gov/pubmed/17200478. Accessed 2013-2-1.

21. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research : what, why, and how? 2009;338 (May).

22. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer; 2008.

23. Wyatt JC, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? BMJ 1995 Dec;311 (7019):1539–1541. http://www.bmj.com/content/311/7019/1539#ref-14. Accessed 2013-2-3.

24. Walker J. Data Mining to Recruit Sick People. Wall Str. J. 2013; http://www.wsj.com/articles/SB10001424052702303722104579240140554518458. Accessed 2015-1-2.

25. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. Biometrics 2005 Mar;61 (1):92–105. http://www.ncbi.nlm.nih.gov/pubmed/15737082. Accessed 2013-2-4.

26. Harrell FEJ, Lee KL, Mark DB. Tutorial in biostatistics Multivariable prognostic models: issues in developing models, evaluating assumptions and adequancy, and measuring and reducing errors. Stat. Med. 1996;15:361–387.

27. Johnstone IM, Titterington DM. Statistical challenges of high-dimensional data. Philos. Trans. A. Math. Phys. Eng. Sci. 2009 Nov;367 (1906):4237–53. http://rsta.royalsocietypublishing.org/content/367/1906/4237. Accessed 2014-12-2.

28. Schrodi SJ, Mukherjee S, Shan Y, Tromp G, Sninsky JJ, Callear AP, Carter TC, Ye Z, Haines JL, Brilliant MH, Crane PK, Smelser DT, Elston RC, Weeks DE. Genetic-based prediction of disease traits: prediction is very difficult, especially about the future. Front. Genet. 2014 Jan;5:162. http://journal.frontiersin.org/Journal/10.3389/fgene.2014.00162/abstract. Accessed 2014-7-13.

29. Wang TJ, Gona P, Larson MG, Tofler GH, Levy D, Newton-Cheh C, Jacques PF, Rifai N, Selhub J, Robins SJ, Benjamin EJ, D'Agostino RB, Vasan RS. Multiple biomarkers for the prediction of first major cardiovascular events and death. N. Engl. J. Med. 2006 Dec;355 (25):2631–9. http://www.ncbi.nlm.nih.gov/pubmed/17182988. Accessed 2013-6-11.

30. Pencina MJ, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat. Med. 2008 Jan;27 (2):157–72; discussion 207–12. http://www.ncbi.nlm.nih.gov/pubmed/17569110. Accessed 2012-10-30.

31. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010 Jan;21 (1):128–38. http://www.ncbi.nlm.nih.gov/pubmed/20010215. Accessed 2012-11-8.

32. Hilden J, Gerds TA. A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. Stat. Med. 2013 Apr; http://www.ncbi.nlm.nih.gov/pubmed/23553436. Accessed 2013-8-20.

33. Lord SR, Sherrington C, Menz HB. Falls in Older People: Risk Factors and Strategies for Prevention. Cambridge University Press; 2001.

34. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk Factors for Falls in Community-dwelling Older People: A Systematic Review and Meta-analysis. Epidemiology 2010;21 (5):658–668.

35. Deandrea S, Bravi F, Turati F, Lucenteforte E, La Vecchia C, Negri E. Risk factors for falls in older people in nursing homes and hospitals. A systematic review and meta-analysis. Arch. Gerontol. Geriatr. 2013;56 (3):407–15. http://www.ncbi.nlm.nih.gov/pubmed/23294998. Accessed 2013-12-19.

36. Leipzig RM, Cumming RG, Tinetti ME. Drugs and falls in older people: a systematic review and meta-analysis: II. Cardiac and analgesic drugs. J. Am. Geriatr. Soc. 1999 Jan;47 (1):40–50. http://www.ncbi.nlm.nih.gov/pubmed/9920228. Accessed 2014-9-2.

37. Leipzig RM, Cumming RG, Tinetti ME. Drugs and falls in older people: a systematic review and meta-analysis: I. Psychotropic drugs. J. Am. Geriatr. Soc. 1999 Jan;47 (1):30–9. http://www.ncbi.nlm.nih.gov/pubmed/9920227. Accessed 2014-9-2.

38. Moreland JD, Richardson JA, Goldsmith CH, Clase CM. Muscle weakness and falls in older adults: a systematic review and meta-analysis. J. Am. Geriatr. Soc. 2004 Jul;52 (7):1121–9. http://www.ncbi.nlm.nih.gov/pubmed/15209650. Accessed 2014-9-2.

39. Stubbs B, Binnekade T, Eggermont L, Sepehry A a, Patchay S, Schofield P. Pain and the risk for falls in community-dwelling older adults: systematic review and meta-analysis. Arch. Phys. Med. Rehabil. 2014 Jan;95 (1):175–187.e9. http://www.ncbi.nlm.nih.gov/pubmed/24036161. Accessed 2014-9-2.

40. Dhital A, Pey T, Stanford MR. Visual loss and falls: a review. Eye (Lond). 2010 Sep;24 (9):1437–46. http://dx.doi.org/10.1038/eye.2010.60. Accessed 2012-11-13.

41. Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the “get-up and go” test. Arch. Phys. Med. Rehabil. 1986 Jun;67 (6):387–9. http://www.ncbi.nlm.nih.gov/pubmed/3487300. Accessed 2014-4-1.

42. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. J. Am. Geriatr. Soc. 1991 Feb;39 (2):142–8. http://www.ncbi.nlm.nih.gov/pubmed/1991946. Accessed 2014-3-27.

43. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can. J. Public Health 1992 Jan;83 Suppl 2:S7–11. http://www.ncbi.nlm.nih.gov/pubmed/1468055. Accessed 2014-12-16.

44. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J. Am. Geriatr. Soc. 1986 Feb;34 (2):119–26. http://www.ncbi.nlm.nih.gov/pubmed/3944402. Accessed 2014-3-25.

45. American Geriatrics Society, British Geriatric Society. Prevention of falls in older persons. AGS/BGS clinical practice guideline. [date unknown]; http://www.medcats.com/FALLS/frameset.htm. .

46. Centers for Disease Control and Prevention. CDC - STEADI - Older Adult Falls - Home and Recreational Safety - Injury Center. [date unknown]; http://www.cdc.gov/homeandrecreationalsafety/Falls/steadi/. Accessed 2015-1-4.

47. Stevens JA, Phelan EA. Development of STEADI: a fall prevention resource for health care providers. Health Promot. Pract. 2013 Sep;14 (5):706–14. http://hpp.sagepub.com/content/early/2012/10/24/1524839912463576.abstract. Accessed 2015-1-6.

48. Centers for Disease Control and Preventio. CDC - Older Adult Falls - Publications - Home and Recreational Safety - Injury Center. [date unknown]; http://www.cdc.gov/HomeandRecreationalSafety/Falls/pubs.html. Accessed 2015-1-5.

49. Centers for Disease Control and Prevention. Algorithm for fall risk assessment & intervention. [date unknown]; http://www.cdc.gov/homeandrecreationalsafety/pdf/steadi/algorithm_fall_risk_assessment.pdf. Accessed 2015-2-21.

50. Centre for Clinical Practice at NICE. Falls. Assessment and prevention of falls in older people. 2013.

51. National Institute for Health and Care Excellence. Falls in older people overview - NICE Pathways. [date unknown]; http://pathways.nice.org.uk/pathways/falls-in-older-people. Accessed 2015-3-3.

52. ProFouND. Profound | Profound Portal. [date unknown]; http://profound.eu.com/. Accessed 2015-1-7.

53. Gates S, Smith LA, Fisher JD, Lamb SE. Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. J. Rehabil. Res. Dev. 2008 Jan;45 (8):1105–16. http://www.ncbi.nlm.nih.gov/pubmed/19235113. Accessed 2012-11-3.

54. Persad CC, Cook S, Giordani B. Assessing falls in the elderly: should we use simple screening tests or a comprehensive fall risk evaluation? Eur. J. Phys. Rehabil. Med. 2010 Jun;46 (2):249–59. http://www.ncbi.nlm.nih.gov/pubmed/20485227. Accessed 2013-2-7.

55. Oliver D, Daly F, Martin FC, McMurdo MET. Risk factors and risk assessment tools for falls in hospital in-patients: a systematic review. Age Ageing 2004 Mar;33 (2):122–30. http://www.ncbi.nlm.nih.gov/pubmed/14960426. Accessed 2013-1-28.

56. Da Costa BR, Rutjes AWS, Mendy A, Freund-Heritage R, Vieira ER. Can falls risk prediction tools correctly identify fall-prone elderly rehabilitation inpatients? A systematic review and meta-analysis. PLoS One 2012 Jan;7 (7):e41061. http://dx.plos.org/10.1371/journal.pone.0041061. Accessed 2013-2-2.

57. Myers H. Hospital fall risk assessment tools: a critique of the literature. Int. J. Nurs. Pract. 2003 Aug;9 (4):223–35. http://www.ncbi.nlm.nih.gov/pubmed/12887374. .

58. Fabre JM, Ellis R, Kosma M, Wood RH. Literature Review Falls Risk Factors and a Compendium of Falls Risk Screening Instruments. J. Geriatr. Phys. Ther. 2010;33 (4):184–197.

59. Scott V, Votova K, Scanlan A, Close J. Multifactorial and functional mobility assessment tools for fall risk among older adults in community , home-support , long-term and acute care settings. Age Ageing 2007;36 (February):130–139.

60. Perell KL, Nelson A, Goldman RL, Luther SL, Prieto-Lewis N, Rubenstein LZ. Fall risk assessment measures: an analytic review. J. Gerontol. A. Biol. Sci. Med. Sci. 2001 Dec;56 (12):M761–6. http://www.ncbi.nlm.nih.gov/pubmed/11723150. .

61. Close JCT, Lord SR. Fall assessment in older people. BMJ 2011 Sep;343 (sep14 1) http://www.bmj.com/content/343/bmj.d5153. Accessed 2012-11-5.

62. Lee J, Geller AI, Strasser DC. Analytical review: focus on fall screening assessments. PM R 2013 Jul;5 (7):609–21. http://www.sciencedirect.com/science/article/pii/S1934148213001652. Accessed 2014-4-10.

63. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. 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 Jan;14 (1):14.

64. Schoene D, Wu SM-S, Mikolaizak AS, Menant JC, Smith ST, Delbaere K, Lord SR. 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 Feb;61 (2):202–8. http://www.ncbi.nlm.nih.gov/pubmed/23350947. Accessed 2013-5-23.

65. Oliver D, Papaioannou A, Giangregorio L, Thabane L, Reizgys K, Foster G. A systematic review and meta-analysis of studies using the STRATIFY tool for prediction of falls in hospital patients: how well does it work? Age Ageing 2008 Nov;37 (6):621–7. http://ageing.oxfordjournals.org/cgi/content/abstract/37/6/621. Accessed 2012-4-2.

66. 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 Jan;13 (1):76. http://www.biomedcentral.com/1471-2296/13/76. Accessed 2013-1-31.

67. Palumbo P, Palmerini L, Mellone S, Cappello A, Chiari L. Towards a standardized protocol for fall risk assessment using wearable sensors. In: Proceedings of the Third Conference of the Italian National Bioengineering Group (GNB2012). 2012 p. 270–271.

68. Howcroft J, Kofman J, Lemaire ED. Review of fall risk assessment in geriatric populations using inertial sensors. J. Neuroeng. Rehabil. 2013 Aug;10 (1):91. http://www.ncbi.nlm.nih.gov/pubmed/23927446. Accessed 2013-8-16.

69. Ejupi A, Lord SR, Delbaere K. New methods for fall risk prediction. Curr. Opin. Clin. Nutr. Metab. Care 2014;17 (5):407–411. http://ovidsp.tx.ovid.com/sp-3.13.1a/ovidweb.cgi?QS2=434f4e1a73d37e8c504983e01e46c99a61d29ae524496461cd0215d59cd93d8162e5751783f43631b22a2767003d84dba94b041cba8dc1477cb304e113950fd7dec9af6d6f528b6502a373164d6835960c63cdb4cbd751d37af1cf48f3fb58cd75108eb0148fcc6aa8d15b503d9d154f1c28ba561206f41797add90b473343e1d0140b2ff9ebfc7ad28acbac6cc431943b0ba5e734021f9402a89bd7a201996cf51946ab0b1e11ed8d70e51c2e5ce094634eaae79b49fbdd2a80ae58191da04ae6c7d1205f8f207073f382f341f8e7d432165d52643052916f3713cf8cbbaf9. Accessed 2014-11-21.

70. Bogle Thorbahn LD, Newton RA. Use of the Berg Balance Test to predict falls in elderly persons. Phys. Ther. 1996 Jun;76 (6):576–83; discussion 584–5. http://www.ncbi.nlm.nih.gov/pubmed/8650273. .

71. Raîche M, Hébert R, Prince F, Corriveau H. Screening older adults at risk of falling with the Tinetti balance scale. Lancet 2000 Sep;356 (9234):1001–2. http://www.ncbi.nlm.nih.gov/pubmed/11041405. Accessed 2013-3-25.

72. Trueblood PR, Hodson-Chennault N, McCubbin A, Yougclarke D. Performance and Impairment-Based Assessments Among Community Dwelling Elderly: Sensitivity and Specificity. J. Geriatr. Phys. Ther. 2001;24 (1):2–6. http://journals.lww.com/jgpt/Abstract/2001/24010/Performance_and_Impairment_Based_Assessments_Among.1.aspx. Accessed 2013-3-25.

73. Faber MJ, Bosscher RJ, van Wieringen PCW. Clinimetric properties of the performance-oriented mobility assessment. Phys. Ther. 2006 Jul;86 (7):944–54. http://ptjournal.apta.org/content/86/7/944.full. Accessed 2012-9-21.

74. Morris R, Harwood RH, Baker R, Sahota O, Armstrong S, Masud T. A comparison of different balance tests in the prediction of falls in older women with vertebral fractures: a cohort study. Age Ageing 2007 Jan;36 (1):78–83. http://www.ncbi.nlm.nih.gov/pubmed/17264139. Accessed 2013-3-11.

75. Oliver D, Britton M, Seed P, Martin FC, Hopper a H. Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. BMJ 1997 Oct;315 (7115):1049–53. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2127684&tool=pmcentrez&rendertype=abstract. .

76. Lamb SE, McCabe C, Becker C, Fried LP, Guralnik JM. The optimal sequence and selection of screening test items to predict fall risk in older disabled women: the Women's Health and Aging Study. J. Gerontol. Med. Sci. 2008;.

77. Delbaere K, Close JCT, Heim J, Sachdev PS, Brodaty H, Slavin MJ, Kochan NA, Lord SR. A multifactorial approach to understanding fall risk in older people. J. Am. Geriatr. Soc. 2010 Sep;58 (9):1679–85. http://www.ncbi.nlm.nih.gov/pubmed/20863327. Accessed 2012-10-11.

78. Marschollek M, Gövercin M, Rust S, Gietzelt M, Schulze M, Wolf K-H, Steinhagen-Thiessen E. Mining geriatric assessment data for in-patient fall prediction models and high-risk subgroups. BMC Med. Inform. Decis. Mak. 2012 Jan;12 (1):19. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3314576&tool=pmcentrez&rendertype=abstract. Accessed 2012-4-29.

79. Lord SR, Menz HB, Tiedemann A. A physiological profile approach to falls risk assessment and prevention. Phys. Ther. 2003 Mar;83 (3):237–52. http://ptjournal.apta.org/content/83/3/237.full. Accessed 2013-6-18.

80. Lee H-C, Chang K-C, Tsauo J-Y, Hung J-W, Huang Y-C, Lin S-I. Effects of a multifactorial fall prevention program on fall incidence and physical function in community-dwelling older adults with risk of falls. Arch. Phys. Med. Rehabil. 2013 Apr;94 (4):606–15, 615.e1. http://www.sciencedirect.com/science/article/pii/S0003999312012014. Accessed 2015-1-16.

81. Delbaere K, Close JCT, Brodaty H, Sachdev P, Lord SR. Determinants of disparities between perceived and physiological risk of falling among elderly people: cohort study. BMJ 2010 Aug;341 (aug18 4):c4165–c4165. http://www.bmj.com/content/341/bmj.c4165?view=long&pmid=20724399. Accessed 2012-10-11.

82. Keogh C, Wallace E, O'Brien KK, Galvin R, Smith SM, Lewis C, Cummins A, Cousins G, Dimitrov BD, Fahey T. Developing an international register of clinical prediction rules for use in primary care: a descriptive analysis. Ann. Fam. Med. 2014 Jul;12 (4):359–66. http://www.annfammed.org/content/12/4/359.long. Accessed 2015-2-25.

83. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann. Intern. Med. 2006 Feb;144 (3):201–9. http://www.ncbi.nlm.nih.gov/pubmed/16461965. Accessed 2014-3-17.

84. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research : application and impact of prognostic models in clinical practice. BMJ 2009;338 (june):1487–1490.

85. Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, Verbakel J, Lakhanpaul M, Fahey T. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). BMC Med. Inform. Decis. Mak. 2011 Jan;11 (1):62. http://www.biomedcentral.com/1472-6947/11/62. Accessed 2015-2-25.

86. Meyer G, Köpke S, Bender R, Mühlhauser I. Predicting the risk of falling--efficacy of a risk assessment tool compared to nurses' judgement: a cluster-randomised controlled trial [ISRCTN37794278]. BMC Geriatr. 2005 Jan;5 (1):14. http://www.biomedcentral.com/1471-2318/5/14. Accessed 2015-2-24.

87. Meyer G, Köpke S, Haastert B, Mühlhauser I. Comparison of a fall risk assessment tool with nurses' judgement alone: a cluster-randomised controlled trial. Age Ageing 2009 Jul;38 (4):417–23. http://www.ncbi.nlm.nih.gov/pubmed/19435759. Accessed 2015-2-24.

88. Rizzo JA, Baker DI, McAvay G, Tinetti ME. The cost-effectiveness of a multifactorial targeted prevention program for falls among community elderly persons. Med. Care 1996 Sep;34 (9):954–69. http://www.ncbi.nlm.nih.gov/pubmed/8792783. Accessed 2015-2-26.

89. Girosi F, Olmsted SS, Keeler E, Hay Burgess DC, Lim Y-W, Aledort JE, Rafael ME, Ricci KA, Boer R, Hilborne L, Derose KP, Shea M V., Beighley CM, Dahl CA, Wasserman J. Developing and interpreting models to improve diagnostics in developing countries. Nature 2006 Nov;.

90. Wu S, Keeler EB, Rubenstein LZ, Maglione MA, Shekelle PG. A cost-effectiveness analysis of a proposed national falls prevention program. Clin. Geriatr. Med. 2010 Nov;26 (4):751–66. http://www.ncbi.nlm.nih.gov/pubmed/20934620. Accessed 2013-8-25.

91. Smith RD, Widiatmoko D. The cost-effectiveness of home assessment and modification to reduce falls in the elderly. Aust. N. Z. J. Public Health 1998 Jun;22 (4):436–40. http://www.ncbi.nlm.nih.gov/pubmed/9659769. Accessed 2013-9-25.

92. Hektoen LF, Aas E, Lurås H. Cost-effectiveness in fall prevention for older women. Scand. J. Public Health 2009 Aug;37 (6):584–9. http://www.ncbi.nlm.nih.gov/pubmed/19666674. Accessed 2013-9-25.

93. Frick KD, Kung JY, Parrish JM, Narrett MJ. Evaluating the cost-effectiveness of fall prevention programs that reduce fall-related hip fractures in older adults. J. Am. Geriatr. Soc. 2010 Jan;58 (1):136–41. http://www.ncbi.nlm.nih.gov/pubmed/20122044. Accessed 2013-8-12.

94. Greenwood M, Yule GU. An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents. J. R. Stat. Soc. 1920;83 (2):255–279.

95. FARSEEING project. [date unknown]; http://farseeingresearch.eu/. .

Chapter 2

The content of this chapter is mainly taken from [1]. We present a probabilistic model to address critical issues about fall prediction through the analysis of the properties of an ideal prognostic tool for falls.

Introduction

Besides assessment tools whose principal aim is to identify key risk factors to target factor-specific interventions [2; 3] — which will not be dealt with in this Chapter— common prognostic tools either produce a continuum score related to the probability of occurrence of one or more falls in a given amount of time [4], or are simple categorical prediction tools that stratify the population according to a dichotomous "low-versus-high risk" logic [5]. Systematic reviews have generally remarked that:

- only few tools have been externally validated [6] (i.e. have been evaluated in a population different from the one employed in the development phase, see [7; 8] for a comprehensive discussion about validation of prognostic models);
- not a single validated tool has so far shown excellent discriminative properties [5; 6; 9];
- heterogeneity of population characteristics and study settings affect the predictive properties of the tools [5; 10; 11].

From the literature in biostatistics and epidemiology it is known that even a '*perfect*' prognostic tool, i.e. a tool that assigns each subject their true probability[1] to develop the outcome of interest [12], cannot reach perfect discrimination [13; 14], and that the upper limit for the area under the receiver operating characteristic (ROC) curve (AUC) depends on the distribution of risk in the population [12; 15]. Nevertheless the clinical literature about falls has never discussed its results in light of these theoretical considerations and some recently-proposed tools incorporating inertial sensors data [16] have shown good but unlikely results. Furthermore, it is known that if a clinical or biological marker is used to predict the time until the development of a given outcome, the ROC curve of the associated prognostic tool is dependent on the censoring time [17]. Finally, the studies that have developed or evaluated fall risk assessment tools have considered as outcome of interest the condition of having fallen either at least once (e.g. [18]) or at least twice during the follow-up (e.g. [19]). How the follow-up length of a prospective study and the definition of

[1] When referring to 'true probability', we adopt the model of probabilities as objective propensity scores.

the outcome of interest impact on the estimated predictive properties of the fall risk tool under evaluation has not been investigated yet.

Objectives

By means of a probabilistic model we aim to investigate the above mentioned theoretical issues about fall prediction. Setting a framework where an ideal prognostic tool for fall risk is evaluated in a population enrolled in a hypothetical prospective clinical study, we aim to derive analytically and evaluate quantitatively its predictive and discriminative performances, and to investigate how these performances are affected by i) the distribution of the fall rate in the population, ii) the follow-up duration, and iii) the definition of faller as single faller or multiple faller.

Methods

Probabilistic model

We assume to evaluate a prognostic tool for falls in an infinite population within a prospective study. All the subjects $\omega_1, \omega_2, \ldots$ of this population are followed over time, from $t = 0$, instant of the baseline, until $t = \tau$, duration of the follow-up.

To each subject ω_i of the population, we associate two random variables, namely N_i and Λ_i. N_i is a random variable accounting for the number of falls that ω_i will experience during the follow-up. Conditional on Λ_i, N_i is assumed independent of any other random variables of the model. With this assumption the number of falls that a subject ω_i will experience during the follow-up depends only on the value of Λ_i. Λ_i is the expected fall rate of ω_i (expressed as the number of falls per year). Then, Λ_i is interpreted as a measure of the proneness to falling of ω_i. We assume the random variables Λ_1, Λ_2,… to be independent and identically distributed according to a distribution F.

Thus, (N_1, Λ_1), (N_2, Λ_2), … are independent and identically distributed (i.i.d.) couples of random variables. We call (N, Λ) one of these i.i.d. couples. We assume that N has a conditional Poisson distribution with mean $\lambda\tau$ given $\Lambda = \lambda$. Its conditional probability mass function is thus:

$$g(n; \lambda\tau) \equiv P(N = n | \Lambda = \lambda) = e^{-\lambda\tau} \frac{(\lambda\tau)^n}{n!} \quad (1)$$

The marginal probability mass function for N is clearly:

$$h(n) \equiv P(N = n) = \int_0^{+\infty} g(n; \lambda\tau) dF(\lambda) \quad (2)$$

Equation (2) says that N follows a mixture of Poisson distributions, with mixing distribution F. Thus fall counts are regarded as arising from a mixture of subjects, each falling according to a Poisson law conditioned on their expected fall rate, with the expected fall rate being distributed according to F. As supported by empirical evidence [20], we hypothesize the marginal distribution for N to be a negative binomial:

$$N \sim NegBin(k, \mu) \quad (3)$$

k and μ being the two real and positive parameters that characterize the distribution. Its probability mass function is:

$$h(n) = \frac{\Gamma(n+k)}{n!\Gamma(k)}\left(\frac{\mu}{\mu+k}\right)^n\left(\frac{k}{\mu+k}\right)^k \quad (4)$$

Here we refer to identifiability as the property to univocally determine the mixing distribution (the distribution F of Λ) given the mixture distribution (the distribution of N) and the conditional distribution, $N|\Lambda = \lambda$. Since continuous mixtures of Poisson distributions are identifiable [21; 22], we deduce that the distribution of Λ, F, is a gamma distribution:

$$\Lambda \sim Gamma(k, \theta) \quad (5)$$

with k, shape parameter of the distribution, being equal to the parameter k of the negative binomial distribution, and ϑ, scale parameter of the distribution, being determined by $\theta = \mu/(k\tau)$. The probability density function of Λ, derivative of F, is:

$$f(\lambda) = \frac{\lambda^{k-1}e^{-\lambda/\theta}}{\Gamma(k)\theta^k} \quad (6)$$

Its mean is $k\theta$ and its variance is $k\theta^2$.

Thus, N conforms to the model introduced by Greenwood and Yule for accident-proneness [23]. Appendix 2 recalls the negative binomial distribution and derives the relation between its parameters and the parameters of the gamma mixing distribution.

We then evaluate the performances of an ideal prognostic tool for falls, r. r is the defined as a function that assigns to each subject ω_i the value of their expected fall rate $\Lambda_i = \lambda_i$, i.e. $r(\omega_i) = \lambda_i$. Thus, according to the definition of a perfect prognostic tool given in the Introduction, r is perfect. The discriminative and predictive performances of this prognostic tool are calculated according to the formulas shown in Table 1 [24]. Following the two alternative approaches usually employed in clinical studies [25], fallers have been defined as those that during the follow-up have fallen either at least once ($\bar{n} = 0$) or at least twice ($\bar{n} = 1$). According to these two definitions, in the following we shall refer to 'prediction of any fall' or 'prediction of multiple falls', respectively.

Table 1. Formulas defining discriminative and predictive performance indices.

Sensitivity	$Se(\lambda_c) = P(\Lambda > \lambda_c \mid N > \bar{n})$
Specificity	$Sp(\lambda_c) = P(\Lambda \leq \lambda_c \mid N \leq \bar{n})$
PPV	$PPV(\lambda_c) = P(N > \bar{n} \mid \Lambda > \lambda_c)$
NPV	$NPV(\lambda_c) = P(N \leq \bar{n} \mid \Lambda \leq \lambda_c)$
Accuracy	$Acc(\lambda_c) = P(N > \bar{n}, \Lambda > \lambda_c) + P(N \leq \bar{n}, \Lambda \leq \lambda_c)$
AUC	$AUC = \int_0^{+\infty} Se(\lambda_c)\frac{dSp}{d\lambda_c}(\lambda_c)d\lambda_c$

λ_c is a given cutoff value on the fall rate.

Parameter estimation and results visualization

We estimated the distribution F for four different populations of community-dwelling and congregate-living older adults (Sydney, Melbourne, New Zealand and Atlanta). The data on fall counts were taken from [20]. Similarly to what already done in [20], for each population we fitted an intercept-only negative binomial regression, estimating the parameters k and μ. Negative binomial regression was done with the function *glm.nb* of the package MASS of the R software [26]. We then derived the parameter k' and θ of the distribution of Λ as $k' = k$ and $\theta = \frac{\mu}{\tau k}$.

MATLAB (R2011a) [27] has been used to plot and visually inspect the analytical formulas of Table 2. The AUC has been calculated via trapezoidal rule for numerical integration with the MATLAB function *trapz* .

Results

Analytic results: performance indices

Under the parametric assumptions presented in Methods, the joint probability of having a fall rate exceeding a given cutoff λ_c and experiencing more than $\bar{n}$ falls is

$$P(\Lambda > \lambda_c,\ N > \bar{n}) = \int_{\lambda_c}^{+\infty} \sum_{n=\bar{n}+1}^{+\infty} g(n;\lambda\tau)\, f(\lambda) d\lambda \qquad (7)$$

whereas the marginal probability of experiencing more than $\bar{n}$ falls is

$$P(N > \bar{n}) = \sum_{n=\bar{n}+1}^{+\infty} h(n) = 1 - \sum_{n=0}^{\bar{n}} h(n). \qquad (8)$$

According to the definition given in Table 1, the sensitivity of the ideal prognostic tool is

$$Se(\lambda_c) = \frac{P(\Lambda>\lambda_c, N>\bar{n})}{P(N>\bar{n})} = \frac{\int_{\lambda_c}^{+\infty} \sum_{n=\bar{n}+1}^{+\infty} g(n;\lambda\tau) f(\lambda) d\lambda}{\sum_{n=\bar{n}+1}^{+\infty} h(n)}. \qquad (9)$$

We define the upper incomplete gamma function as $\gamma_U(x,k) \coloneqq \frac{1}{\Gamma(k)} \int_x^{+\infty} s^{k-1} e^{-s} ds$, Γ being the gamma function, i.e. $\Gamma(k) = \int_0^{+\infty} s^{k-1} e^{-s} ds$. Recalling the relation $\mu = \theta\tau k$ and the property $\sum_{n=0}^{+\infty} g(n;\lambda\tau) = 1$, the computations for the case $\bar{n} = 0$ (prediction of any fall) proceeds as follows

$$Se(\lambda_c) = \frac{\int_{\lambda_c}^{+\infty} \sum_{n=1}^{+\infty} g(n;\lambda\tau) f(\lambda) d\lambda}{\sum_{n=1}^{+\infty} h(n)} = \frac{\int_{\lambda_c}^{+\infty} [1-g(0;\lambda\tau)] f(\lambda) d\lambda}{1-h(0)} =$$
$$\frac{\frac{1}{\Gamma(k)} \int_{\lambda_c}^{+\infty} \theta^{-k} \lambda^{k-1} e^{-\lambda/\theta} d\lambda - \frac{1}{\Gamma(k)} \int_{\lambda_c}^{+\infty} e^{-\lambda\tau} \theta^{-k} \lambda^{k-1} e^{-\lambda/\theta} d\lambda}{1-(1+\theta\tau)^{-k}} = \frac{\frac{1}{\Gamma(k)} \int_{\lambda_c/\theta}^{+\infty} s^{k-1} e^{-s} ds - \frac{(1+\theta\tau)^{-k}}{\Gamma(k)} \int_{\lambda_c/\theta(1+\theta\tau)}^{+\infty} s^{k-1} e^{-s} ds}{1-(1+\theta\tau)^{-k}} =$$
$$\frac{\gamma_U(\lambda_c/\theta,k) - (1+\theta\tau)^{-k} \gamma_U(\lambda_c/\theta(1+\theta\tau),k)}{1-(1+\theta\tau)^{-k}}$$

(10)

Other formulas for the discriminative and predictive performances of the ideal prognostic tool for falls can be derived similarly. Their expressions, reported in Table 2, have been obtained for both cases of prediction of any fall ($\bar{n} = 0$), and prediction of multiple falls ($\bar{n} = 1$).

Table 2. Discriminative and predictive performance indices of the ideal prognostic tool for falls, for the two definitions of faller.

	$\boldsymbol{Fallers} := \{\boldsymbol{\omega_i}: \boldsymbol{N_i} = \boldsymbol{n_i} > \bar{\boldsymbol{n}},\ \bar{\boldsymbol{n}} = \boldsymbol{0}\}$
Sensitivity	$\frac{\gamma_U(\lambda_c/\theta\,,k) - (1+\theta\tau)^{-k}\gamma_U(\lambda_c/\theta\,(1+\theta\tau),k)}{1-(1+\theta\tau)^{-k}}$
Specificity	$\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k)$
PPV	$\frac{\gamma_U(\lambda_c/\theta\,,k) - (1+\theta\tau)^{-k}\gamma_U(\lambda_c/\theta\,(1+\theta\tau),k)}{\gamma_U(\lambda_c/\theta\,,k)}$
NPV	$\frac{(1+\theta\tau)^{-k}\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k)}{\gamma_L(\lambda_c/\theta\,,k)}$
Accuracy	$\gamma_U(\lambda_c/\theta\,,k) + (1+\theta\tau)^{-k}[-1+2\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k)]$
	$\boldsymbol{Fallers} := \{\boldsymbol{\omega_i}: \boldsymbol{N_i} = \boldsymbol{n_i} > \bar{\boldsymbol{n}},\ \bar{\boldsymbol{n}} = \boldsymbol{1}\}$
Sensitivity	$\frac{\gamma_U(\lambda_c/\theta\,,k) - (1+\theta\tau)^{-k}\gamma_U(\lambda_c/\theta\,(1+\theta\tau),k) - \mu(1+\theta\tau)^{-k-1}\gamma_U(\lambda_c/\theta\,(1+\theta\tau),k+1)}{1-(1+\theta\tau)^{-k}-\mu(1+\theta\tau)^{-k-1}}$
Specificity	$\frac{(1+\theta\tau)^{-k}\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k) + \mu(1+\theta\tau)^{-k-1}\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k+1)}{(1+\theta\tau)^{-k}+\mu(1+\theta\tau)^{-k-1}}$
PPV	$\frac{\gamma_U(\lambda_c/\theta\,,k) - (1+\theta\tau)^{-k}\gamma_U(\lambda_c/\theta\,(1+\theta\tau),k) - \mu(1+\theta\tau)^{-k-1}\gamma_U(\lambda_c/\theta\,(1+\theta\tau),k+1)}{\gamma_U(\lambda_c/\theta\,,k)}$
NPV	$\frac{(1+\theta\tau)^{-k}\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k) + \mu(1+\theta\tau)^{-k-1}\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k+1)}{\gamma_L(\lambda_c/\theta\,,k)}$
Accuracy	$\gamma_U(\lambda_c/\theta\,,k) + (1+\theta\tau)^{-k}[-1+2\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k)]$ $+\mu(1+\theta\tau)^{-k-1}[-1+2\gamma_L(\lambda_c/\theta\,(1+\theta\tau),k+1)]$
Notations: $\gamma_L(x,k) = \frac{1}{\Gamma(k)}\int_0^x t^{k-1}e^{-t}dt$; $\gamma_U(x,k) = 1-\gamma_L(x,k)$; $\Gamma(k) = \int_0^{+\infty} t^{k-1}e^{-t}dt$	

Analytic results: accuracy maximization

Following the definition given in Table 1, the accuracy can be computed as:

$$Acc(\lambda_c) = \sum_{n=\bar{n}+1}^{+\infty}\int_{\lambda_c}^{+\infty} g(n;\lambda\tau)f(\lambda)d\lambda + \sum_{n=0}^{\bar{n}}\int_0^{\lambda_c} g(n;\lambda\tau)f(\lambda)d\lambda. \qquad (11)$$

In order to look for a cutoff that maximizes the accuracy, we set its derivative to be zero (necessary condition).

$$\frac{dAcc}{d\lambda_c} = -\sum_{n=\bar{n}+1}^{+\infty} g(n;\lambda_c\tau)f(\lambda_c) + \sum_{n=0}^{\bar{n}} g(n;\lambda_c\tau)f(\lambda_c)$$

$$= f(\lambda_c)\left[-\sum_{n=\bar{n}+1}^{+\infty} g(n;\lambda_c\tau) + \sum_{n=0}^{\bar{n}} g(n;\lambda_c\tau)\right]$$

$$= f(\lambda_c)\left[-1 + 2\sum_{n=0}^{\bar{n}} g(n;\lambda_c\tau)\right] = 0 \qquad (12)$$

$\lambda_{c,max}$ that satisfies equation (11) is then determined by

$$\sum_{n=0}^{\bar{n}} g\left(n;\lambda_{c,max}\tau\right) = \frac{1}{2} \qquad (13)$$

For the case $\bar{n} = 0$, $\lambda_{c,max}$ is simply $\log(2)/\tau$. It is easy to show that the second derivative of Acc in $\lambda_c = \lambda_{c,max}$ given in equation (12) is negative for every $\bar{n}$ nonnegative integer. Thus $\lambda_{c,max}$ maximizes the accuracy. This cutoff does not depend on the distribution of the fall rate in the population, but just on $\bar{n}$ and τ.

Quantitative results

The parameters estimates for the distribution of Λ in the four populations are the following: Sydney *k*=0.47, ϑ=1.71; Melbourne *k*=1.14, ϑ=0.81; New Zealand *k*=0.80, ϑ=1.10; Atlanta *k*=1.70, ϑ=0.77. These have been estimated as explained in Methods and considering that in these studies the follow-up duration τ is respectively 0.46 years, 1 year, 1 year, and 0.92 years.

Figure 1 and Table 3 report the discriminative and predictive performances of the ideal tool for different values of the parameters k and θ, a follow-up duration of 1 year and $\bar{n} = 0$. The AUC and the maximum accuracy are non-linear and non-monotonic functions of the parameters of the fall rate distribution. For the four populations the AUC ranges from 0.80 to 0.89. Assuming a cutoff value λ_c=log(2) falls/year, the sensitivity of the tool ranges from 0.71 to 0.84, the specificity from 0.57 to 0.87, the accuracy from 0.74 to 0.81, the positive predictive value (PPV) from 0.73 to 0.77, the negative predictive value (NPV) from 0.68 to 0.84.

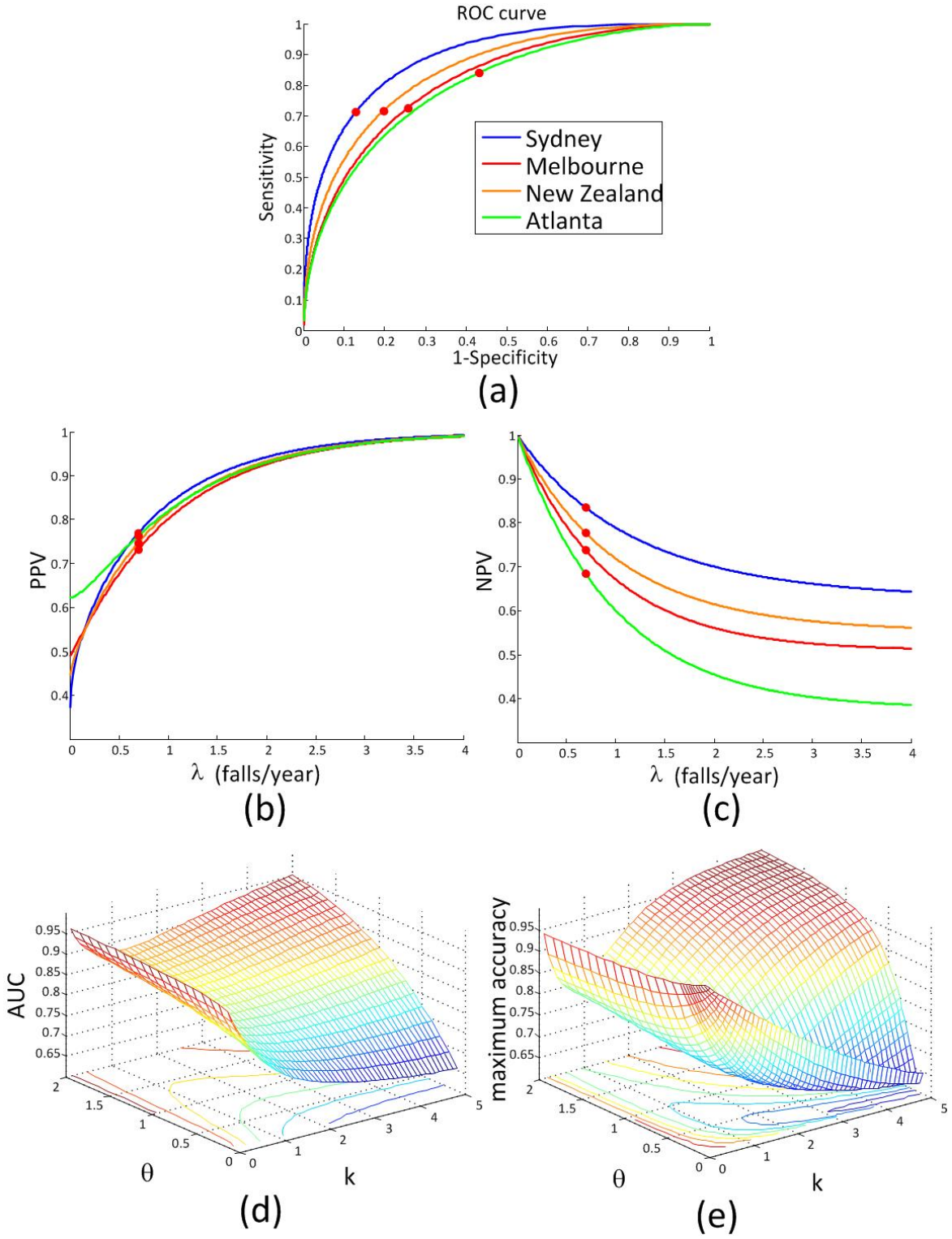

Figure 1. Performances of the ideal prognostic tool for different parameters (*k* and ϑ) of the population fall rate distributions, follow-up length τ=1 year, $\overline{n} = 0$ (prediction on any fall). *k* and ϑ parameterize the distribution of the fall rate Λ so that its mean is $k\theta$ and its variance is $k\theta^2$. In panels (a)-(c): the performances of the ideal tool have been evaluated using the estimated shape and scale parameters *k* and ϑ of the four populations: Sydney, Melbourne, New Zealand, Atlanta. The red dots mark the points corresponding to a cutoff value λ_c=log(2) falls/year. Panels (d), (e): AUC and maximum accuracy (accuracy calculated for $\lambda = \lambda_{c,max}$) as functions of *k* and ϑ.

Table 3. Discriminative and predictive performance of the ideal prognostic tool for falls evaluated on four populations ($\lambda_c = \log(2)\ fall/year, \tau = 1\ year, \bar{n} = 0$)

Population	Sensitivity	Specificity	PPV	NPV	Accuracy	AUC
Sydney	0.71	0.87	0.77	0.84	0.81	0.89
Melbourne	0.73	0.74	0.73	0.74	0.74	0.82
New Zealand	0.72	0.80	0.75	0.78	0.76	0.85
Atlanta	0.84	0.57	0.76	0.68	0.74	0.80

The sensitivity of the performance indices with respect to the follow-up duration is shown in Figure 2, for the representative parameters extracted from the Sydney population. The AUC is only slightly affected, ranging from 0.84 at 8 weeks to 0.91 at 2 years. The PPV (NPV) is a monotonically increasing (decreasing) function of the follow-up duration. It ranges from 0.22 (0.98) at 8 weeks, to 0.88 (0.77) at 2 years.

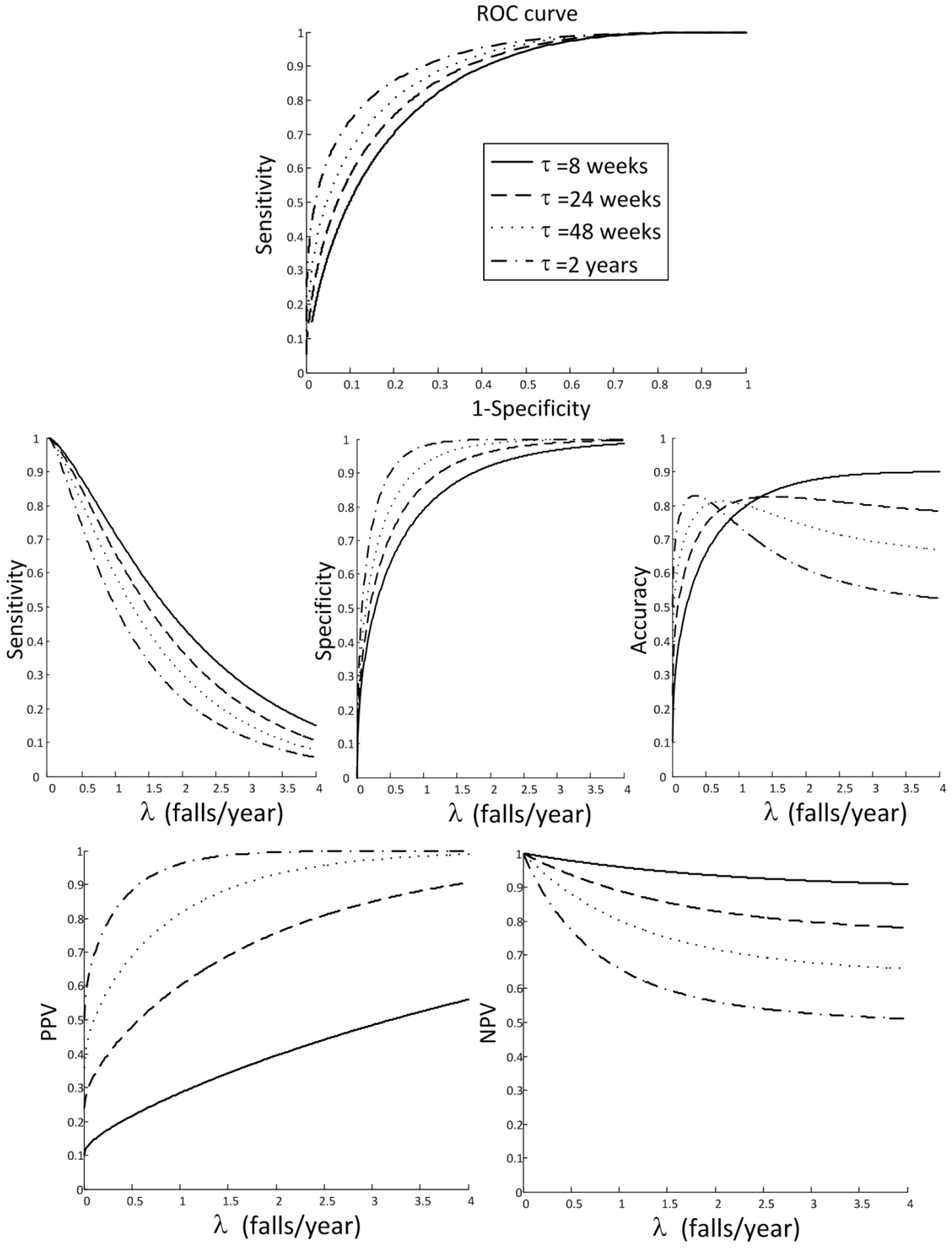

Figure 2. Sensitivity of the discriminative and predictive performance indices to the follow-up duration τ. Parameters are estimated from the Sydney population (k=0.47, ϑ=1.71), $\bar{n} = 0$ (prediction on any fall).

Figure 3 shows the discriminative and predictive performances of the tool when prediction is made on any fall or multiple falls ($\bar{n} = 0$ or $\bar{n} = 1$), for the representative parameters extracted from the Sydney population. The AUC's for the four populations and the two predictions are: (Sydney, $\bar{n} = 0$) = 0.89, (Sydney, $\bar{n} = 1$) = 0.92; (Melbourne, $\bar{n} = 0$) = 0.82, (Melbourne, $\bar{n} = 1$) = 0.86; (New Zealand, $\bar{n} = 0$) = 0.85, (New Zealand, $\bar{n} = 1$) = 0.89; (Atlanta, $\bar{n} = 0$) = 0.80, (Atlanta, $\bar{n} = 1$) = 0.83.

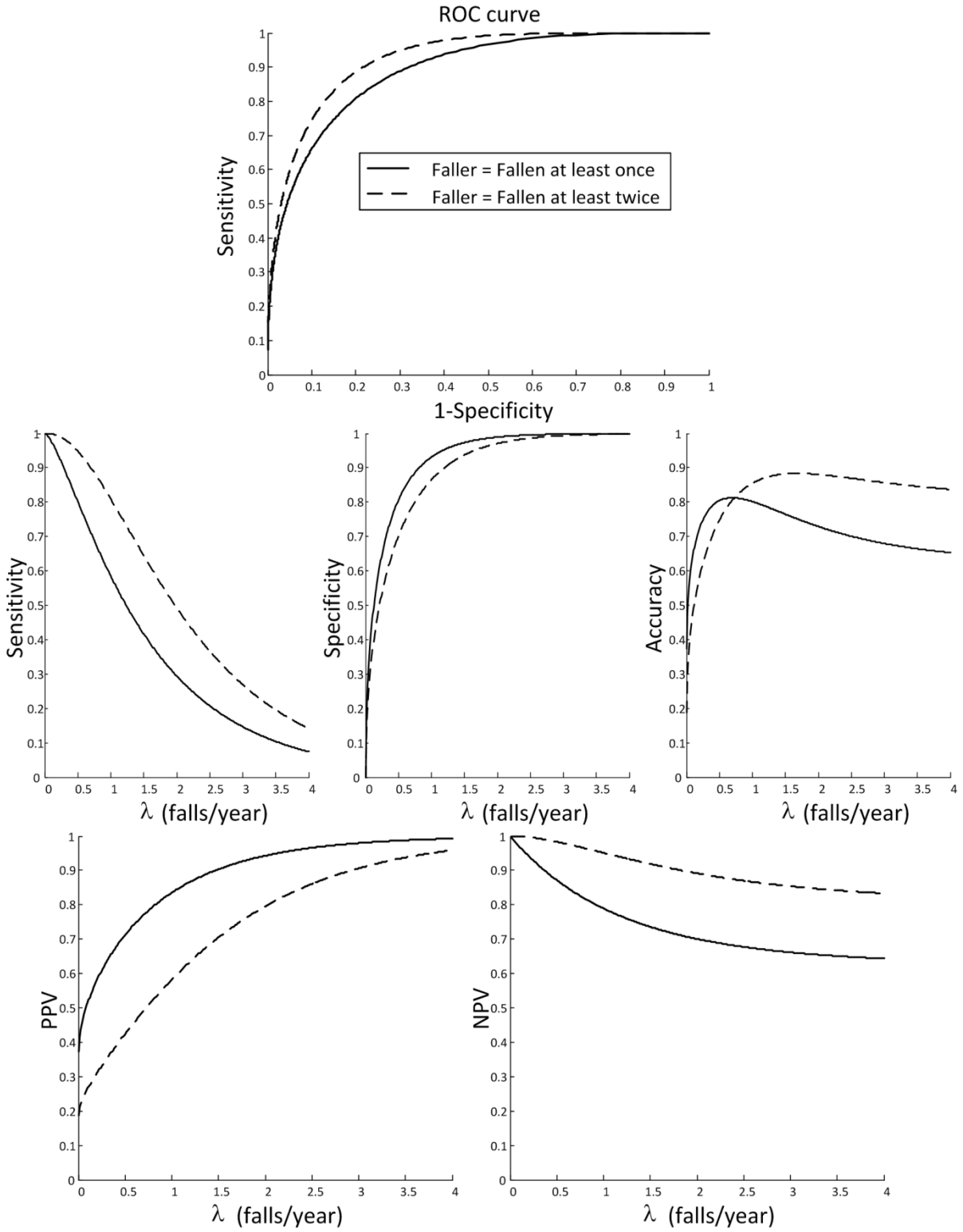

Figure 3. Performance indices of the ideal prognostic tool employing the two alternative definitions of fallers, as those fallen at least once (solid line, $\bar{n} = 0$, prediction on any fall) or at least twice (dashed line, $\bar{n} = 1$, prediction on multiple falls) during the follow-up. Parameters are estimated from the Sydney population (*k*=0.47, ϑ=1.71); follow-up length $\tau = 1$ year.

Discussion

In this Chapter the predictive and discriminative performances of an ideal prognostic tool for falls have been evaluated by means of a probabilistic model. The indices considered for the evaluation have been the sensitivity, specificity, AUC, accuracy, and positive and negative predictive values. Although other metrics could have been considered (e.g. the Brier score and fractional reduction in entropy [12]), these have been chosen because they are by far the most commonly employed.

Having thus obtained the performances of a perfect prognostic tool for falls allows a critical assessment of some results that can be found when evaluating real prognostic tools. While it is known that, despite considerable research efforts, externally validated clinical tools still have modest performances [5; 9; 10], it is not infrequent that newly-developed tools come up with excellent but unlikely results. Having at hand some indicative reference values for the upper bounds of indices quantifying the goodness of the prediction is methodologically advisable and can suggest warnings against over-optimism.

The problem of over-optimism, often affecting newly-developed prognostic tools, has already been highlighted in the literature [7; 28]. A well-studied example is the STRATIFY (St Thomas Risk Assessment Tool in Elderly Inpatients). Without going into the details of the development and validation of this tool, it is worth mentioning that after being tested in several cohorts of older in-patients, a review [9] concluded that its prognostic performances are sensibly lower than previously reported by the first studies that led to its publication [29]. Another example could be represented by some recent sensor-based tools that have shown perfect accuracy [16].

One of the factors that may influence the reproducibility of prognostic tools (i.e. their capacity to keep their performance on subjects not included in the dataset used for original development, but similar for characteristics) is a low ratio between number of cases (number of fallers in our case) and number of candidate predicting variables available at the development stage (see the number of events per variable, EPV, discussed in [7]). This factor is critical in a context, like the development of sensor-based prognostic tools for falls, where (as yet) there is high availability of candidate variables. In this case, using statistical techniques that properly manage the high dimensionality of the problem (leveraging the so called 'bias versus variance tradeoff' [30]) and performing appropriately internal validation (e.g. cross-validation) are crucial. Furthermore, it is noteworthy that over-optimism may arise in the literature via publication bias even applying correct procedures of model fitting and validation. When the sample size is small, the estimation of the performance indices is subject to high variability, and studies with better results will be more likely published. Finally, we point out that when the sample size is small, because of the variability on the estimated performance, even an imperfect prognostic tool for falls can outperform the limits here calculated. Reporting confidence intervals for the estimated parameters should hence be recommended in real applications. In our study the perfect tool was evaluated in the entire population, i.e. no sampling process has been modeled.

Among the analytic results, we have proven that the cutoff $\lambda_{c,max}$ that maximizes the accuracy is independent of the population over which the prognostic model is evaluated. This result still holds if the accuracy is modified assigning different weights to true positives and true negatives (the two addends at the right side of equation (10)). Instead, it may not hold for other quantities (e.g. it does not hold for the Youden index). Clinical and economic considerations should lead to key choices for fall prevention strategies, as choices on the frequency of the assessment and the definition of faller (in terms of our notation, of τ and $\bar{n}$). Once these are made, our finding legitimates the practice of choosing a cutoff for a particular fall risk scale and applying it on different populations.

The populations in the Sydney, Melbourne and New Zealand studies are composed of community-dwellers aged 60, 70 and 80 years or more respectively. The population in the Atlanta study is composed of congregate-living, transitionally frail older adults aged 70 years or more. The diversity among these populations is reflected in the estimated parameters. As recalled in Methods, the mean of the fall rate distribution F is given by the product $k\theta$. As expected, this quantity is highest for the Atlanta population (1.31 falls/year) and is lowest for Sydney (0.80 falls/year). In turn, this diversity is responsible of the heterogeneity in the values of the performance indices shown in Table 3 and Figure 1. As shown in Table 3, for a given cutoff, specificity is the parameter that varies most among the four populations, whereas PPV is quite stable. The AUC's that we obtain are much higher than those found on validated clinical tools documented in the literature (see e.g. [10]). Thus, as expected, these traditional tools are far from providing a perfect probabilistic risk assessment.

Consistently over the four populations (results shown in Figure 2 only for the Sydney population), the AUC increases with the length of the follow-up as the net effect of the increase in specificity and decrease in sensitivity. The increase in specificity for a given cutoff can be explained in terms of its components: the proportion of true negatives (TN, subjects with fall rate less than the cutoff and never fallen during the follow-up) and the proportion of false positives (FP, subjects with fall rate higher than the cutoff but never fallen during the follow-up). As the duration of the follow-up increases, more falls occur and both TN and FP decrease. However, as FP have higher expected fall rate, their decrease is quicker and this determines a net increase in specificity. For similar reasons, true positives (subjects with fall rate higher than the cutoff and fallen during the follow-up) and false negatives (subjects with fall rate less than the cutoff and fallen during the follow-up) are responsible for a decrease in sensitivity.

Consistently over the four populations, the AUC is slightly higher when the prediction is made on multiple falls (definition of faller with $\bar{n} = 1$) than on any fall ($\bar{n} = 0$), as the net effect of an increase in sensitivity and a decrease in specificity. Such results may indicate that one same tool is likely to show better discrimination when the prediction is made on multiple falls. This is consistent with what was found and commented in [31], although our findings about the predictability of multiple falls have been reached from a different perspective, without considering any knowledge other than the estimated distribution of the fall rate in the population. Indeed an accurate and accepted definition of who should be classified as a faller is still missing, and a matter of discussion in the literature. Lord and colleagues have proposed to define a faller as one fallen at least twice during the follow-up to filter out 'occasional' falls [2], i.e. with the objective to possibly identify falls which were due only to substantial and persistent risk factors.

Pointing at the distribution of fall rate in the population, the length of the follow-up, and the definition of faller as potential sources of heterogeneity for the reported performances of fall risk prognostic tools, this study links to the work of Haines et al [28], that explained part of the variability found in the literature in terms of differences in study design. Furthermore, in the present study we have showed how the effects of these factors on the discriminative and predictive performances of the tools rely on non-linear relations that standard models for meta-analysis cannot address.

All the results here obtained are valid within the hypotheses stated in Methods and the assumption that the expected fall rate of each subject is constant over time. However, the choice of a Poisson distribution for the conditional number of falls, given the fall rate, accommodates both the scenarios of time-constant and time-variable expected fall rate of each subject in the population, provided that the change in the expected fall rate is independent of the occurrence of a past fall. In particular, calling $N_i(t)$ the random process representing the number of falls from baseline to time t of subject ω_i, the time-constant fall rate

scenario is equivalent to assuming $N_i(t)$ as a homogeneous Poisson process with intensity λ_i, whereas the time-varying fall rate scenario is equivalent to assuming $N_i(t)$ as an inhomogeneous Poisson process with intensity function $\widehat{\lambda_i}(t)$, λ_i being its mean from baseline to time τ: $\lambda_i = \frac{1}{\tau}\int_0^\tau \widehat{\lambda_i}(u)du$. In this second scenario λ_i is clearly dependent on the length of the follow-up. The sensitivity analysis with respect to the length of the follow-up, shown in Figure 2, is no more valid in this second scenario. Therefore, its results should be reconsidered if the expected fall rate of the subjects is believed to undergo substantial changes during the follow-up. In homogeneous and inhomogeneous Poisson processes the occurrence of an event is independent of the occurrence of any other. If this hypothesis is not valid and the change in the expected fall rate is supposed to be driven by the occurrence of a previous fall, other models should be considered (e.g. pure birth process) [32; 33]. However, it is worth noting that, given only the fall counts in a given time period, the identifiability among alternative models is not guaranteed [34]. Thus far, all the clinical tools have followed the approach of providing a unique score for the proneness to falling, without discerning for scenarios of subject-specific time-varying fall rate during the follow-up, nor has clinical epidemiology provided sufficient descriptive evidence for this kind of scenarios. Our choice to give a main focus to the case of constant fall rate has to be considered in this light and for the sake of simplicity.

The gamma distribution for Λ was deduced from the hypothesis that the marginal distribution of the number of falls is of negative binomial type and from the identifiability of continuous mixtures of Poisson distributions [21; 22]. The hypothesis for this marginal distribution is supported from the empirical evidence shown in [20]. A gamma distribution for the fall rate has already been considered for negative binomial regression [35; 36]. The theoretical results about the identifiability of Poisson distributions makes the problem of estimating F well posed. However, given a finite number of observations over the N_i's, we cannot exclude that other distributions may fit equally well the data.

Conclusions

We have proposed the model of an ideal prognostic tool for falls which operates within a population according to the Greenwood and Yule scheme for accident proneness. We have derived analytically the performance indices of such perfect prognostic tool. We have estimated the parameters of the fall rate distribution of four different populations observed in different epidemiological studies, and we have then obtained quantitative evaluation of the analytical formulas. In the four considered populations, the AUC of the perfect tool, predicting any fall over a follow-up of one year, was estimated to range between 0.80 and 0.89.

We have showed that the performance indices of the perfect prognostic tool can be estimated solely from falls counts and can be useful reference values for future works introducing new fall risk assessment tools. The analytical results give also an indication about how to choose a cutoff that maximizes the accuracy or any other weighted function of true positive and true negative rates. The maximum accuracy when prediction is made on any fall for a follow-up length τ is attained with a cutoff of $\log(2)/\tau$.

The model has allowed us to identify, analyze and quantify the effect of major factors that account for the high heterogeneity of results observed in the literature: i) the fall rate distribution over the population, ii) the length of the follow-up, and iii) the definition of faller as single faller or multiple faller. Because of the different fall rate distributions, specificity was found to have remarkable variations, varying over the four considered populations from 0.57 to 0.87. Predicting on multiple falls was found to have an effect on the AUC in terms of an increase of about 0.04 with respect to prediction on any fall.

References

1. Palumbo P, Palmerini L, Chiari L. A probabilistic model to investigate the properties of prognostic tools for falls. Methods Inf. Med. 2014;53 (6).

2. Lord SR, Menz HB, Tiedemann A. A physiological profile approach to falls risk assessment and prevention. Phys. Ther. 2003 Mar;83 (3):237–52. http://ptjournal.apta.org/content/83/3/237.full. Accessed 2013-6-18.

3. Tiedemann A, Lord SR, Sherrington C. The development and validation of a brief performance-based fall risk assessment tool for use in primary care. J. Gerontol. A. Biol. Sci. Med. Sci. 2010 Aug;65 (8):896–903. http://www.ncbi.nlm.nih.gov/pubmed/20522529. Accessed 2013-5-23.

4. Hernandez D, Rose DJ. Predicting which older adults will or will not fall using the Fullerton Advanced Balance scale. Arch. Phys. Med. Rehabil. 2008 Dec;89 (12):2309–15. http://www.ncbi.nlm.nih.gov/pubmed/18976981. Accessed 2012-7-16.

5. Oliver D, Papaioannou A, Giangregorio L, Thabane L, Reizgys K, Foster G. A systematic review and meta-analysis of studies using the STRATIFY tool for prediction of falls in hospital patients: how well does it work? Age Ageing 2008 Nov;37 (6):621–7. http://ageing.oxfordjournals.org/cgi/content/abstract/37/6/621. Accessed 2012-4-2.

6. Scott V, Votova K, Scanlan A, Close J. Multifactorial and functional mobility assessment tools for fall risk among older adults in community , home-support , long-term and acute care settings. Age Ageing 2007;36 (February):130–139.

7. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat. Med. 2000 Feb;19 (4):453–473. http://www.ncbi.nlm.nih.gov/pubmed/10694730. .

8. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research : validating a prognostic model. BMJ 2009;338 (june):1432–1435.

9. 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 Jan;13 (1):76. http://www.biomedcentral.com/1471-2296/13/76. Accessed 2013-1-31.

10. Gates S, Smith LA, Fisher JD, Lamb SE. Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. J. Rehabil. Res. Dev. 2008 Jan;45 (8):1105–16. http://www.ncbi.nlm.nih.gov/pubmed/19235113. Accessed 2012-11-3.

11. Muir SW, Berg K, Chesworth B, Klar N, Speechley M. Quantifying the magnitude of risk for balance impairment on falls in community-dwelling older adults: a systematic review and meta-analysis. J. Clin. Epidemiol. 2010 Apr;63 (4):389–406. http://dx.doi.org/10.1016/j.jclinepi.2009.06.010. Accessed 2012-3-25.

12. Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. Biostatistics 2005 Apr;6 (2):227–39. http://biostatistics.oxfordjournals.org/content/6/2/227.abstract. Accessed 2013-1-28.

13. Diamond G. What price perfection? Calibration and discrimination of clinical prediction models. J. Clin. Epidemiol. 1992;45 (1):85–89.

14. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clin. Chem. 2008 Jan;54 (1):17–23. http://www.clinchem.org/content/54/1/17.full. Accessed 2012-11-8.

15. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007 Feb;115 (7):928–35. http://circ.ahajournals.org/content/115/7/928.long. Accessed 2012-10-26.

16. Howcroft J, Kofman J, Lemaire ED. Review of fall risk assessment in geriatric populations using inertial sensors. J. Neuroeng. Rehabil. 2013 Aug;10 (1):91. http://www.ncbi.nlm.nih.gov/pubmed/23927446. Accessed 2013-8-16.

17. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 2000 Jun;56 (2):337–44. http://www.ncbi.nlm.nih.gov/pubmed/10877287. .

18. Marschollek M, Rehwald a, Wolf KH, Gietzelt M, Nemitz G, Meyer Zu Schwabedissen H, Haux R. Sensor-based fall risk assessment--an expert "to go". Methods Inf. Med. 2011 Jan;50 (5):420–6. http://www.ncbi.nlm.nih.gov/pubmed/21206963. Accessed 2012-11-27.

19. Stalenhoef PA, Diederiks JPM, Knottnerus JA, Kester ADM, Crebolder HFJM. A risk model for the prediction of recurrent falls in community-dwelling elderly: a prospective cohort study. J. Clin. Epidemiol. 2002 Nov;55 (11):1088–94. http://www.ncbi.nlm.nih.gov/pubmed/12507672. .

20. Ullah S, Finch CF, Day L. Statistical modelling for falls count data. Accid. Anal. Prev. 2010 Mar;42 (2):384–92. http://dx.doi.org/10.1016/j.aap.2009.08.018. Accessed 2012-5-23.

21. Feller W. On a General Class of "Contagious" Distributions. Ann. Math. Stat. 1943;14 (4):389–400.

22. Johnson NL, Kotz S, Kemp AW. Univariate discrete distributions. Second edi. New York: Wiley; 1992.

23. Greenwood M, Yule GU. An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents. J. R. Stat. Soc. 1920;83 (2):255–279.

24. Metz CE. Basic Principles of ROC Analysis. Semin. Nucl. Med. 1978;VIII (4):283–298.

25. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk Factors for Falls in Community-dwelling Older People: A Systematic Review and Meta-analysis. Epidemiology 2010;21 (5):658–668.

26. R Core Team, R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing. 2013; http://www.r-project.org. .

27. MATLAB and Simulink for Technical Computing - MathWorks Italia. [date unknown]; http://www.mathworks.it/. .

28. Haines TP, Hill K, Walsh W, Osborne R. Design-Related Bias in Hospital Fall Risk Screening Tool Predictive Accuracy Evaluations: Systematic Review and Meta-Analysis. J. Gerontol. Med. Sci. 2007;62A (6):664–672.

29. Oliver D, Britton M, Seed P, Martin FC, Hopper a H. Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. BMJ 1997 Oct;315 (7115):1049–53. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2127684&tool=pmcentrez&rendertype=abstract. .

30. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data Mining, Inference, and Prediction. Second edi. Springer; 2009.

31. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. A prospective study. JAMA 1989 May;261 (18):2663–8. http://www.ncbi.nlm.nih.gov/pubmed/2709546. Accessed 2013-7-6.

32. Papoulis A. Probability, Random Variables, and Stochastic Processes. 3rd editio. New York: McGraw-Hill; 1991.

33. Hsu HP. Theory and Problems of Probability, Random Variables, and Random Processes. italian ed. McGraw-Hill; 2011.

34. Xekalaki E. The univariate generalized Waring distribution in relation to accident theory: proneness, spells or contagion? Biometrics 1983;39 (4):887–895. http://www.jstor.org/stable/10.2307/2531324?Search=yes&resultItemClick=true&searchText=A&searchText=Class&searchText=of&searchText=Non-Identifiable&searchText=Stochastic&searchText=Models&searchUri=/action/doBasicSearch?Query=A+Class+of+Non-Identifiable+Stochastic+Models&Search=Search&gw=jtx&prq=%28au%3A%28Cane%29+AND+ti%3A%28A+Class+of+Non-Identifiable+Stochastic+Models%29%29&hp=25&acc=on&a. Accessed 2014-4-4.

35. Navarro A, Ancizu I. Analyzing the occurrence of falls and its risk factors: some considerations. Prev. Med. (Baltim). 2009 Mar;48 (3):298–302. http://www.ncbi.nlm.nih.gov/pubmed/19168092. Accessed 2012-7-22.

36. Robertson MC, Campbell AJ, Herbison P. Statistical analysis of efficacy in falls prevention trials. J. Gerontol. A. Biol. Sci. Med. Sci. 2005 Apr;60 (4):530–4. http://www.ncbi.nlm.nih.gov/pubmed/15933397. .

Chapter 3

Introduction

In this Chapter we present the development and validation of a fall risk assessment tool, named FRAT-up [1–3]. This tool has been developed within the framework of the EU project FARSEEING in collaboration with the group of Artificial Intelligence of the University of Bologna.

All its parameters are derived exclusively from the literature. Its validation is carried out using data from three different epidemiological studies: 'Invecchiare nel Chianti' (InCHIANTI) [4; 5], the ActiFE-Ulm study [6–8], and the English Longitudinal Study of Ageing (ELSA) [9; 10].

Architecture

The architecture of FRAT-up differentiates between risk estimators and risk factors. A sketch is provided in Figure 1. Risk estimators are quantitative information from clinical tests or questionnaires that is used to derive the exposure of a given subject to a particular risk factor. FRAT-up receives risk estimators in input. It then derives exposure to risk factors from estimators, and uses exposure to risk factors to calculate the risk of falling at least once in the time span of one year. The addition of an intermediate level between input and output (i.e. the risk factors between risk estimators and risk of falling) gives the tool the capability for adapting to different inputs after changing only the mapping between the first two layers, i.e. it gives the tool the capability of accepting in input new, different risk estimators, given that a function is known for deriving the risk factors from the new risk estimators.

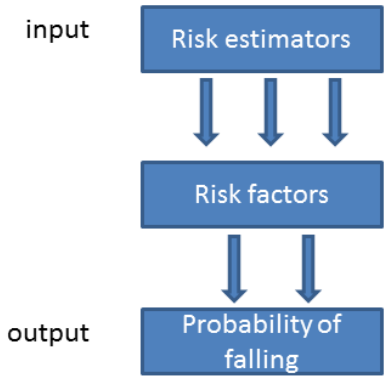

Figure 1. Sketch of FRAT-up architecture.

The list of the risk factors and their types (whether they are continuous or dichotomous) is taken from the review and meta-analysis by Deandrea et al [11]. It is reported below in Table 1. Some additional information is available in Appendix 3.

Table 1. List of risk factors for falls considered by FRAT-up.

Domain	Name	Type
Socio-demographic	age	scalar
	female sex	dichotomous
	living alone	dichotomous
	history of falls	dichotomous
	physical activity limitation	dichotomous
	physical disability	dichotomous
	instrumental disability	dichotomous
	walking aid use	dichotomous
Medical and psychological	cognition impairment	dichotomous
	depression	dichotomous
	history of stroke	dichotomous
	urinary incontinence	dichotomous
	rheumatic disease	dichotomous
	dizziness and vertigo	dichotomous
	diabetes	dichotomous
	comorbidity	scalar
	poor self-perceived health status	dichotomous
	pain	dichotomous
	fear of falling	dichotomous
	Parkinson disease	dichotomous
Medications	number of medications	scalar
	use of sedatives	dichotomous
	use of antihypertensives	dichotomous
	use of antiepileptics	dichotomous
Mobility and sensory	gait problems	dichotomous
	vision impairment	dichotomous
	hearing impairment	dichotomous

Details about the conversion rules from risk estimators to risk factors for three different datasets can be found in Appendix 3-InCHIANTI, Appendix 3-ActiFE, Appendix 3-ELSA.

Model formulation and parameter derivation

Base model

Let $E_0, E_1, ..., E_n$ be $n+1$ dichotomous random variables with values in $\{0;1\}$, and $E = (E_0, E_1, ..., E_n)$. We say that the i^{th} risk factor is present if $E_i = 1$. Let $d_0, d_1, ..., d_n$ be $n+1$ events. We assume the following conditional independence relations:

$$d_i \mid E_i \perp d_j, E_j \forall j \neq i \qquad (1).$$

We call d_i a fall event specific to risk factor E_i (intuitively, we think at d_i as a fall caused by E_i).

Assumptions from Equation (1) can be phrased saying that risk factor-specific falls are mutually independent conditional on their associated risk factor.

We define the event d as the union of the factor-specific events d_i's:

$$d := \bigcup_{i=0}^{n} d_i$$

i.e. d is verified if at least one of the d_i's is. This is an assumption of causal independence where the "causes" $E_0, E_1, ..., E_n$ contribute independently to the probability of the effect d; for a complete formal definition see [12]. d is the presence of at least one fall event during a given time span, while E is an observation of the risk factor exposures of a subject before the time span.

The conditional probability of d given E can then be calculated as follows, by De Morgan laws and assumptions in Equation (1):

$$P(d \mid E) = P\left(\bigcup_{i=0}^{n} d_i \mid E\right) = 1 - P\left(\overline{\bigcup_{i=0}^{n} d_i} \mid E\right) =$$

$$= 1 - P\left(\bigcap_{i=0}^{n} \overline{d_i} \mid E\right) =$$

$$= 1 - \prod_{i=0}^{n} P\left(\overline{d_i} \mid E_i\right) =$$

$$= 1 - \prod_{i=0}^{n} \left[1 - P(d_i \mid E_i)\right] \qquad (2).$$

This function, modeling the probability of an event given a set of possible causes, is known as noisy-OR gate [13].

We assume that

$$P(d_i \mid E_i) = \begin{cases} 0 & if\ E_i = 0 \\ C_i & if\ E_i = 1 \end{cases} \quad (3)$$

C_i is the contribution to the probability of the effect d given by the exposure to the risk factor E_i; a method to assign values to the contributions C_i is introduced in the following. Using Equation (3), Equation (2) becomes

$$P(d \mid E) = 1 - \prod_{i=0}^{n} (1 - E_i C_i) \quad (4)$$

Since we want to model a minimum probability of falling that is applied even in the absence of any observation-specific exposures, we assign $P(E_0 = 1) = 1$. C_0 is the risk that is present in this case.

Parameter derivation

The contributions C_i of the exposures have been derived from the odds ratios (OR) computed in the meta-analysis by Deandrea et al [11]. In the following we present the assumptions and the calculations.

The OR relative to risk factor E_i, with $i = 1,...,n$, is defined as:

$$OR_i := \frac{P(d \mid E_0 = 1, E_i = 1)}{1 - P(d \mid E_0 = 1, E_i = 1)} \frac{1 - P(d \mid E_0 = 1, E_i = 0)}{P(d \mid E_0 = 1, E_i = 0)}$$

Note that the condition $E_0 = 1$ is always true and is highlighted above just for convenience.

We assume that OR_i may be approximated as

$$OR_i = \frac{P(d \mid E_0 = 1, E_i = 1, E_{j \neq 0,i} = 0)}{1 - P(d \mid E_0 = 1, E_i = 1, E_{j \neq 0,i} = 0)} \frac{1 - C_0}{C_0} \quad (5).$$

Informally, this assumption states that the OR computed on the whole population is similar to the OR computed restricting the population to subjects having at most one exposure. This is obviously true in models where each subject has at most one exposure; otherwise there is a difference in the two values.

Given assumptions in Equations (1) and (3),

$$P(d \mid E_0 = 1, E_i = 1, E_{j \neq 0,i} = 0) =$$

$$= P(d_0 \vee d_i \mid E_0 = 1, E_i = 1, E_{j \neq 0,i} = 0) =$$

$$= P(d_0 \mid E_0 = 1) + P(d_i \mid E_i = 1) - P(d_0 \mid E_0 = 1) P(d_i \mid E_{i=1}) =$$

$$C_0 + C_i - C_0 C_i \quad (6)$$

Substituting Equation (6) in Equation (5) and solving for C_i we finally get

$$C_i = C_0 \frac{OR_i - 1}{1 - C_0 + C_0 OR_i} \tag{7}$$

and substituting it in Equation (4)

$$P(d \mid E) = 1 - \prod_{i=0}^{n} \left(1 - E_i C_0 \frac{OR_i - 1}{1 - C_0 + C_0 OR_i} \right).$$

C_0 was calculated by leaving it as a free parameter and then learning it with an equation-solving algorithm. In particular, we used the bisection method, imposing the reported number of total falls from [14].

General model

The model presented above handles only dichotomous risk factors and it requires that we know for every risk factor if it is present or not. Here we present the formulas for a more general case that handles risk factors with more levels, and missing information on risk factor exposure.

Let $E_o, \dots, E_n$ be the list of risk factors. Risk factor E_i can take levels $0,1, \dots, r_i$. P_{ij} is the probability for a given subject that E_i takes level j. $prev_{ij}$ is the probability in the population for E_i to take the level j. Clearly, the followings hold: $\sum_{j=0}^{r_i} P_{ij} = 1, \sum_{j=0}^{r_i} prev_{ij} = 1$.

Let us consider the case of E_i dichotomous risk factor, i.e. $r_i = 1$. If we know that the subject is exposed to risk factor E_i, then $P_{i0} = 0$ and $P_{i1} = 1$. If we know that the subject is not exposed to E_i, then $P_{i0} = 1$ and $P_{i1} = 0$. If information about the exposure of the subject to E_i is missing, then we state $P_{i0} = prev_{i0} = 1 - prev_{i1}$ and $P_{i1} = prev_{i1}$. Similarly it holds for continuous risk factors. If we know that E_i takes the value j, then $P_{ik} = \delta_{kj}$ (δ_{kj} is Kronecker delta). If we do not know the value that E_i takes, then $P_{ij} = prev_{ij}$.

The calculation of the terms P_{ij}'s is a bit more difficult when we consider E_i being the risk factor comorbidity. Let $P_{com,j}$ be these terms, P_{com} be the vector $[P_{com,0}, P_{com,1}, \dots, P_{com,11}]$, and $c_1, c_2, \dots, c_{11}$ be the list of 11 dichotomous risk factors that contribute to risk factor comorbidity (see Appendix 3). Then, $P_{com} = [1 - P_{c1,1},\ P_{c1,1}] * [1 - P_{c2,1},\ P_{c2,1}] * \dots * [1 - P_{c11,1},\ P_{c11,1}]$, where '$*$' denotes a convolution product.

Finally, Equation (4) generalizes to

$$P(d \mid E) = 1 - \prod_{i=0}^{n} \sum_{j=0}^{r_i} P_{ij} (1 - C_i)^j$$

Validation

Once developed, most tools never receive external validation, i.e. they are not tested on a sample from a population that has not been used for the development of the model itself. This happens because of the time and financial cost of collecting new data. Yet, external validation is essential to know the performance of a tool.

The flexibility of FRAT-up expressed by its architecture and its ability to seamlessly deal with missing values allowed us to validate it on three large datasets of three epidemiological studies. In the following, we present methods and results of this validation.

Harmonization and description of variables

Retrospective harmonization is the process of deriving common variables from different existing datasets. It allows the utilization of data coming from different sources within one combined analysis [15].

We call 'target variables' the variables that are wanted as result of the harmonization process. We call 'source variables' (or 'assessment items') the variables that are native of each dataset and that are used to construct the target variables. We distinguish between predictor target variables (all the risk factors considered by FRAT-up, see Table 1) and an outcome target variables (the variable that is the object of prediction, i.e. occurrence of any fall in a given time span after the assessment).

For each dataset, processing algorithms were developed and applied —whenever possible and whenever needed— to construct the target variables from the source variables. The harmonization process was led being blinded to possible influences of the different choices of the harmonization process itself on the performance of any predictive model.

The algorithms to construct the target variables from the source variables of each dataset are given in Appendix 3-InCHIANTI, Appendix 3-ActiFE, and Appendix 3-ELSA. On the ELSA dataset it was considered impossible to construct the variables "Self-perceived health status", "Sedatives", "Antihypertensives", and "Antiepileptics".

Characteristics of the populations

The designs and rationales of the studies ActiFE-Ulm, InCHIANTI, and ELSA are documented in [4; 6; 9].

The description that is here offered of the ActiFE population is relative to the subjects observed during the first wave of the study. The description of the InCHIANTI population is relative to the subjects observed during the first wave of the study and aged 65 years or more. The description of the ELSA population is relative to the subjects observed during the second wave of the study, aged 65 years or more, and that completed the nurse visit. The description of the ELSA population has been carried out using weights, furnished along the dataset, and intended to adjust for differential nonresponse rates to the nurse visit.

Table 2 gives a description of the distributions of the harmonized variables in the three datasets before and after excluding those subjects on which there is no information on the outcome variable ('Prospective falls (yes/no)').

Table 2. Characteristics of the three populations before and after excluding subjects on which there is no information on the outcome variables.

	ActiFE		InCHIANTI		ELSA	
	Before exclusion	After exclusion	Before exclusion	After exclusion	Before exclusion	After exclusion
Number of subjects	1506	1416	1155	892	4056	3303
Age (years): mean (sd)	75.62 (6.59)	75.53 (6.55)	75.44 (7.63)	74.17 (6.91)	75.03 (7.71)	74.56 (7.31)
Sex (women)	0.43	0.43	0.57	0.56	0.56	0.57
History of falls (yes/no)	0.35	0.35	0.23	0.21	0.35	0.34
History of falls (number):	0.92	0.87	0.42	0.35	1.07	0.91

mean (sd)	(8.32)	(8.40)	(1.04)	(0.89)	(8.05)	(4.69)
Prospective falls (yes/no)	0.32	0.32	0.24	0.24	0.34	0.34
Prospective falls (number): mean (sd)	0.63 (1.70)	0.63 (1.70)	0.42 (0.97)	0.42 (0.97)	0.96 (5.21)	0.96 (5.21)
Living alone	0.25	0.24	0.19	0.18	0.34	0.34
Walking aid use	0.02	0.01	0.13	0.08	0.11	0.09
Urinary incontinence	0.38	0.38	0.37	0.35	0.17	0.17
Diabetes	0.14	0.13	0.13	0.13	0.11	0.11
Parkinson disease	0.02	0.02	0.01	0.01	0.01	0.01
Arthritis or rheumatism	0.49	0.49	0.29	0.30	0.45	0.45
Cognition impairment	0.01	0.01	0.17	0.11	0.01	0.01
History of stroke	0.05	0.05	0.05	0.05	0.08	0.07
Depression	0.11	0.11	0.18	0.18	0.11	0.10
Poor self-perceived health status	0.17	0.16	0.08	0.07	NA	NA
Pain	0.59	0.59	0.87	0.87	0.45	0.43
Physical disability	0.04	0.03	0.10	0.05	0.21	0.19
Instrumental disability	0.13	0.12	0.28	0.22	0.17	0.14
Fear of falling	0.11	0.10	0.38	0.37	0.09	0.07
Dizziness	0.40	0.40	0.32	0.34	0.24	0.22
Vision impairment	0.85	0.85	0.55	0.51	0.27	0.25
Hearing impairment	0.25	0.25	0.28	0.27	0.28	0.27
Number of medications: mean (sd)	3.67 (2.94)	3.62 (2.91)	2.34 (2.08)	2.25 (2.04)	0.71 (0.97)	0.69 (0.95)
Use of antihypertensives	0.57	0.56	0.43	0.41	NA	NA
Use of sedatives	0.01	0.01	0.06	0.06	NA	NA
Use of antiepileptics	0.02	0.02	0.02	0.01	NA	NA
Physical activity limitations	0.14	0.13	0.24	0.19	0.10	0.08
Grip strength (Kg): mean (sd)	32.10 (11.13)	32.18 (11.09)	27.59 (11.72)	28.47 (11.78)	25.89 (10.23)	26.38 (10.17)
Gait problems	0.21	0.20	0.24	0.19	0.34	0.32
Gait speed (m/s): mean (sd)	0.97 (0.29)	0.98 (0.28)	1.00 (0.27)	1.02 (0.26)	0.83 (0.26)	0.85 (0.25)
SPPB_BT score: mean (sd)	3.69 (0.81)	3.72 (0.76)	3.22 (1.29)	3.39 (1.13)	3.16 (1.32)	3.27 (1.24)
SPPB_GST score: mean (sd)	3.62 (0.89)	3.65 (0.84)	3.51 (1.02)	3.66 (0.82)	3.41 (0.94)	3.47 (0.89)
SPPB_CST score: mean (sd)	3.17 (1.16)	3.20 (1.13)	2.96 (1.35)	3.11 (1.23)	2.30 (1.49)	2.40 (1.45)
SPPB score: mean (sd)	10.49 (2.35)	10.58 (2.21)	9.72 (3.29)	10.18 (2.78)	9.24 (2.83)	9.46 (2.67)

Performance of FRAT-up

FRAT-up was evaluated on the three populations. It was applied on the harmonized datasets without any previous imputation on missing values, taking advantage of its ability to deal with them. The receiving operating characteristic (ROC) curves were computed with the R package ROCR [16]. 95 % confidence intervals (CI) for the area under the ROC curves (AUC) were computed from 2000 bootstrap replicates with the R package pROC [17]. The Brier score (BS) was computed and decomposed in three terms as proposed

in [18]. The three terms, namely reliability (REL), generalized resolution (GRES), and uncertainty (UNC), are so that BS=REL-GRES+UNC. The BS decomposition and the calibration plots were obtained after dividing the samples in deciles according to the FRAT-up risk score.

The ROC curves are shown in Figure 2. The calibration plots are shown in Figure 3. The BS and its components are shown in Table 3. The AUC on ActiFE is 0.567 (95% CI 0.535-0.599), on InCHIANTI is 0.644 (95% CI 0.601-0.689), on ELSA is 0.704 (95% CI 0.685-0.723). Similarly, the generalized resolution is lowest for ActiFE, higher for InCHIANTI and highest for ELSA. The miscalibration on InCHIANTI (REL=0.01) is explained by over-estimation of the risk, as evident from the calibration plot.

More analyses were done in order to get more insights on these results. These analyses are presented in the next section.

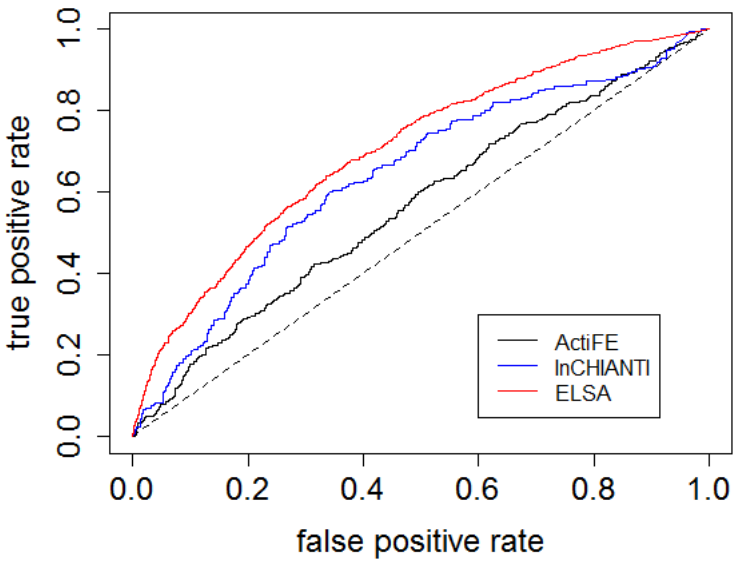

Figure 2. Receiving operating characteristic (ROC) curves of FRAT-up applied on the three populations.

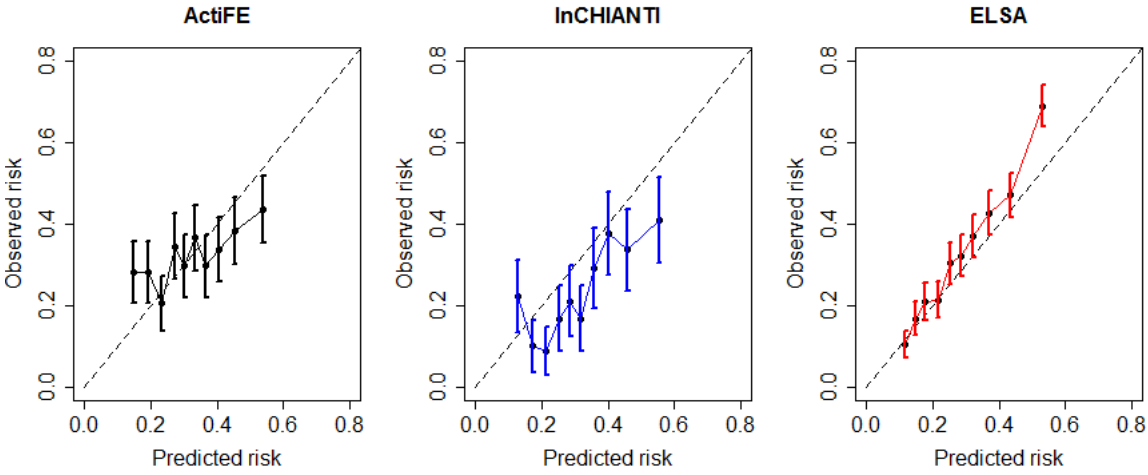

Figure 3. Calibration plots of FRAT-up applied on the three populations.

Table 3. Brier score and its components for FRAT-up on the three populations.

	Brier score	Reliability	Generalized resolution	Uncertainty
ActiFE	0.2201179	0.005819622	0.004530492	0.2188288
InCHIANTI	0.1803304	0.01044918	0.0114689	0.1813501
ELSA	0.1977321	0.003844301	0.02670226	0.2205901

More insights

Trained models

The ActiFE dataset was randomly divided in two disjoint sets, namely a training set and a test set, containing respectively about two thirds and one third of the observations. One of the imputed copies of the training set was used to fit a logistic ridge regression. This regression model was then used to calculate the risk score on the test set of the ActiFE and on the InCHIANTI and ELSA datasets. In particular, the regression model was applied on each imputed copy, obtaining 11 risk scores for each subject. These 11 scores were then averaged to obtain a unique risk score for each subject. This random split of the ActiFE dataset as well as the model fitting and model testing were repeated 20 times. The procedure was then repeated so that the training was in turn made also on InCHIANTI and ELSA.

The results in terms of AUC are shown in Figure 4. Similarly to what observed for FRAT-up, the models perform best on ELSA, worse on InCHIANTI and worst on ActiFE, regardless of the dataset they were trained on.

This likely indicates that the causes of the heterogeneity in model performance obtained on the three datasets have to be sought among the differences of the three datasets rather than be regarded as coming from idiosyncrasies of FRAT-up.

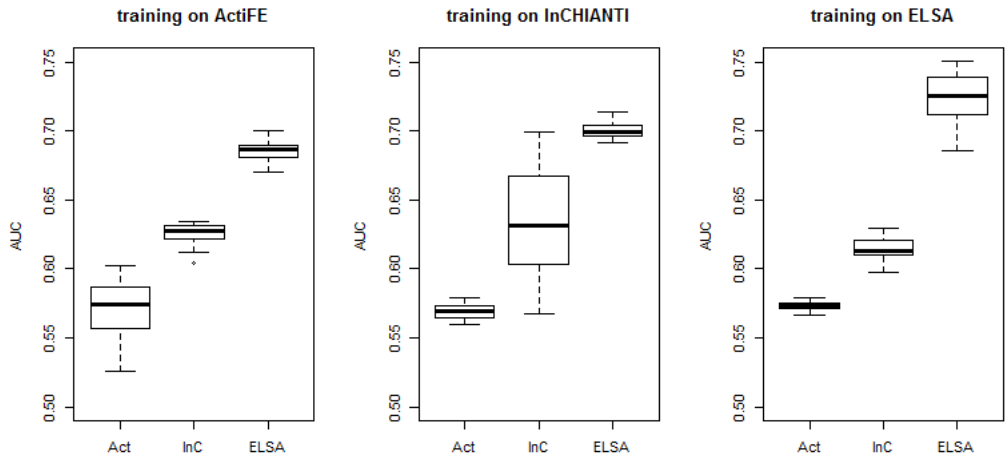

Figure 4. Boxplots for the AUCs of the trained models. The three panels contain results relative to models fitted respectively on ActiFE, InCHIANTI, and ELSA. When a given dataset was used for training, the AUC was computed only on the testing subset. Each boxplot is representative of the AUCs of 20 models, fitted on 20 random splits of the dataset currently used for training.

Fall calendar vs history of falls at follow-up

It is known that data about history of falls during anamnesis can be affected by recall bias [19].

From the ActiFE study, information about falls that occurred after the baseline assessment of the risk factors is available from both fall calendars and as history of falls recalled at the first follow-up. Falls recorded in the calendar are relative to the 12 months after the baseline assessment. Falls recalled at follow-up are relative to the 12 months before the follow-up itself. As the follow-up was carried out about 36 months after the baseline assessment, different time spans are covered from the two information collection methods.

We applied the FRAT-up on covariate data from the baseline assessment of ActiFE and tested whether the performances of the model change substantially if used to predict falls as collected in the fall calendar and at the follow-up. Subjects that did not report the fall calendar, that were lost to follow-up, or that died before it, were labeled as 'not available'.

Table 4 gives the number of subjects that have reported at least one fall according to the two collection methods. Figure 5 shows ROC curves and calibration plots.

The AUC on falls from the calendar is 0.566 (95 % C.I. 53.58-59.78). The AUC on falls from the follow-up is 0.6258 (95 % C.I. 0.5880-0.6670). This increase in AUC reduces the difference with what attained in InCHIANTI and ELSA.

Table 4. Number of fallers and non-fallers according to information from fall calendars and collected at the follow-up.

		Follow-up		
		Non fallers	Fallers	Not available
Fall calendar	Non fallers	376	140	442
	Fallers	161	127	170
	Not available	14	10	66

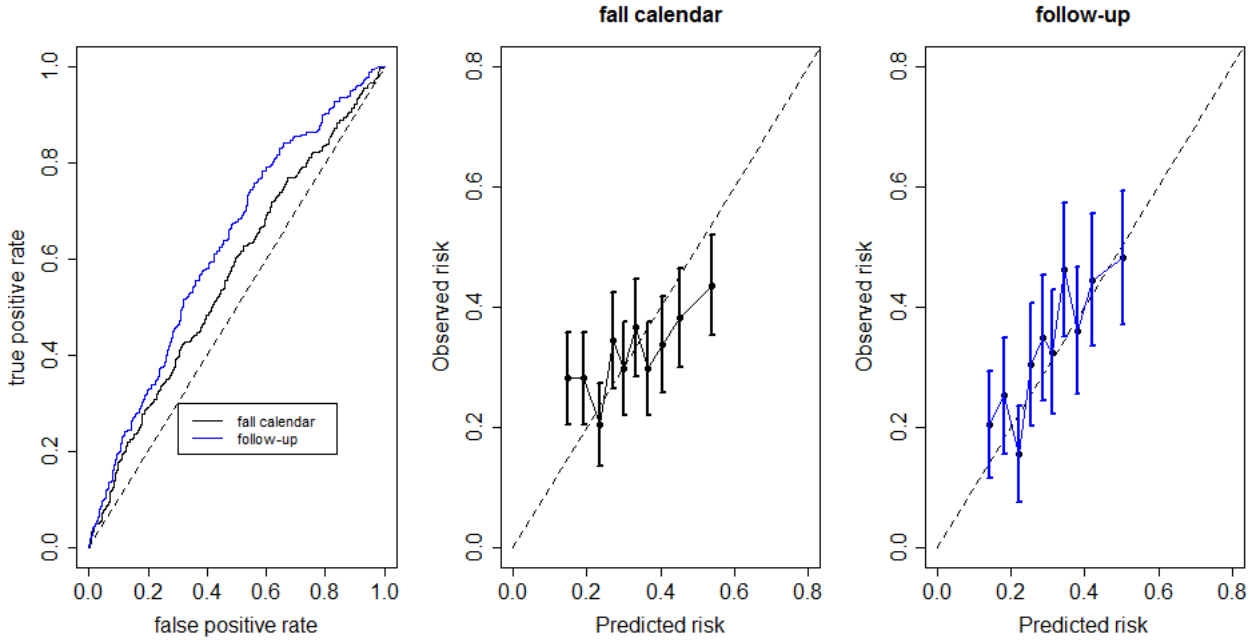

Figure 5. ROC curve (left panel) and calibration plots (central and right panel) for FRAT-up applied to the baseline covariates of ActiFE to predict falls reported in the fall calendar and at the follow-up.

Discussions

Among FRAT-up declared aims there is the characteristic of being flexible enough to allow the use of different estimators for each risk factor [2]. The presumption that the tool would operate properly once different estimators are available in input comes from the way the parameters were calculated. In particular, they were not derived from training on any specific dataset but from the meta-analysis by Deandrea[11], where the effect of each risk factor on falls was derived from heterogeneous studies, employing different estimators for each risk factor. In this Chapter we have presented FRAT-up and applied it on data from three epidemiological studies about ageing, employing different data collection methods.

The results obtained on the three datasets are substantially heterogeneous. This heterogeneity may be imputed to differences among cohorts and differences in data collection methods. For example, ELSA and InCHIANTI aim to be representative of the whole older population, whereas ActiFE has some exclusion criteria (e.g. having a sever deficit in cognition, vision or hearing) and has a higher participation rate from male individuals. Besides differences in risk factors assessment (e.g. three different questionnaires to assess disabilities in activities of daily living), great influence could have been played by different methods to assess falls [20]. In the ActiFE study occurrence of falls has been ascertained by prospective falls calendars, whereas in ELSA and InCHIANTI information about falls is available only through questions about history of falls. The analysis presented in section “Fall calendar vs history of falls at follow-up” may suggest that this difference could explain part of the heterogeneity. Furthermore, questions about history of falls were relative to time intervals of different length (about two years for ELSA, one for ActiFE and InCHIANTI) and starting at different time points after the assessment (immediately after the assessment for ELSA, about two years later for InCHIANTI and ActiFE).

In addition to these sources of heterogeneity, we have to consider the criticalities of the process of retrospective harmonization. Indeed, whether harmonization for a given target variable was considered possible or not, and the choice of processing algorithms to derive it from source variables have been up to our subjective evaluation. Some groups have mitigated this degree of subjectivity with use of consensus methodologies [21]. Other possible approaches could be based on integration of expert opinions with more quantitative tools. For example, the Principal Component Analysis presented in Appendix 3 has highlighted differences in the ELSA population along the second principal component. The variables contributing most to this component could be submitted to a panel of experts for possible revision of their harmonization process. Other quantitative approaches could rely on small pilot populations where different assessment items relative to one same target variable are observed on the same subjects. The datasets thus obtained would allow to calculate correlations among target variables obtained with different processing algorithms, thus having a quantitative validation of the quality of the harmonization.

Despite the ambition to be flexible with respect to the use of different clinical tests to estimate each risk factor [2], FRAT-up web application gives precise guidance about how to enter information about most risk factors (e.g. Mini Mental State Examination is required for assessing cognitive impairment) [22]. The criticalities discussed about the harmonization process and the yet-unexplained heterogeneity of the results obtained suggest that more study is needed before having a validated, fully-flexible tool.

Heterogeneity in the results is a common finding in reviews of studies validating fall risk tools [23–25]. Nowadays, these reviews are available only for few traditional tools and have been published after 20-30 years after the publication of the tool they refer to (see e.g. paper[26]). Our flexible approach allowed us to collect evidence about the performance of FRAT-up on big-size populations using negligible time and financial resources.

References

1. Cattelani L, Chesani F, Palumbo P, Palmerini L, Bandinelli S, Becker C, Chiari L. FRAT-Up, a Rule-Based System Evaluating Fall Risk in the Elderly. In: 2014 IEEE 27th International Symposium on Computer-Based Medical Systems. IEEE; 2014 p. 38–41. http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=6881844. Accessed 2014-9-16.

2. Cattelani L, Palumbo P, Palmerini L, Bandinelli S, Becker C, Chesani F, Chiari L. FRAT-up, a web-based fall risk assessment tool for elderly people living in the community. J. Med. Internet Res. 2015;17 (2).

3. FRAT-Up. [date unknown]; http://ffrat.farseeingresearch.eu/. Accessed 2015-1-17.

4. Ferrucci L, Bandinelli S, Benvenuti E, Di Iorio A, Macchi C, Harris TB, Guralnik JM. 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 Dec;48 (12):1618–25. http://www.ncbi.nlm.nih.gov/pubmed/11129752. Accessed 2013-3-11.

5. InCHIANTI. InCHIANTI Study door. [date unknown]; http://inchiantistudy.net/. Accessed 2013-7-6.

6. Denkinger MD, Franke S, Rapp K, Weinmayr G, Duran-Tauleria E, Nikolaus T, Peter R. 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 Jan;10 (1):50. http://www.biomedcentral.com/1471-2318/10/50. Accessed 2013-12-11.

7. Klenk J, Büchele G, Rapp K, Franke S, Peter R. Walking on sunshine: effect of weather conditions on physical activity in older people. J. Epidemiol. Community Health 2012 May;66 (5):474–6. http://www.ncbi.nlm.nih.gov/pubmed/21325149. Accessed 2015-3-14.

8. Klenk J, Rapp K, Denkinger MD, Nagel G, Nikolaus T, Peter R, Koenig W, Böhm BO, Rothenbacher D. Seasonality of vitamin D status in older people in Southern Germany: implications for assessment. Age Ageing 2013 May;42 (3):404–8. http://www.ncbi.nlm.nih.gov/pubmed/23542723. Accessed 2015-3-14.

9. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort profile: the English Longitudinal Study of Ageing. Int. J. Epidemiol. 2013 Dec;42 (6):1640–8. http://www.ncbi.nlm.nih.gov/pubmed/23143611. Accessed 2014-2-10.

10. Marmot M, Oldfield Z, Clemens S, Blake M, Phelps A, Nazroo J, Steptoe A, Rogers N, Banks J. English Longitudinal Study of Ageing: Waves 0-5, 1998-2011. 2013;.

11. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk Factors for Falls in Community-dwelling Older People: A Systematic Review and Meta-analysis. Epidemiology 2010;21 (5):658–668.

12. Heckerman D, Breese JSJ. A New Look at Causal Independence. Proc. Tenth Int. Conf. Uncertain. Artif. Intell. 1994;286–292.

13. Good I. A Causal Calculus. Br. J. Philos. Sci. 1961;11 (44):305–318.

14. World Health Organization Department of Ageing and Life Course. WHO global report on falls prevention in older age. Geneva, Switzerland: World Health Organization; 2008.

15. Fortier I, Doiron D, Little J, Ferretti V, L'Heureux F, Stolk RP, Knoppers BM, Hudson TJ, Burton PR. Is rigorous retrospective harmonization possible? Application of the DataSHaPER approach across 53 large studies. Int. J. Epidemiol. 2011 Oct;40 (5):1314–28.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3204208&tool=pmcentrez&rendertype=abstract. Accessed 2015-3-5.

16. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005;21 (20):7881. http://rocr.bioinf.mpi-sb.mpg.de. .

17. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011 Jan;12 (1):77. http://www.biomedcentral.com/1471-2105/12/77. Accessed 2013-2-10.

18. Stephenson DB, Coelho C a. S, Jolliffe IT. Two Extra Components in the Brier Score Decomposition. Weather Forecast. 2008 Aug;23 (4):752–757. http://journals.ametsoc.org/doi/abs/10.1175/2007WAF2006116.1. Accessed 2014-8-10.

19. Cummings SR, Nevitt MC, Kidd S. Forgetting falls. The limited accuracy of recall of falls in the elderly. J. Am. Geriatr. Soc. 1988 Jul;36 (7):613–6. http://www.ncbi.nlm.nih.gov/pubmed/3385114. Accessed 2015-3-4.

20. Rapp K, Freiberger E, Todd C, Klenk J, Becker C, Denkinger M, Scheidt-Nave C, Fuchs J. Fall incidence in Germany: results of two population-based studies, and comparison of retrospective and prospective falls data collection methods. BMC Geriatr. 2014 Jan;14:105. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4179843&tool=pmcentrez&rendertype=abstract. Accessed 2014-11-18.

21. Fortier I, Burton PR, Robson PJ, Ferretti V, Little J, L'Heureux F, Deschênes M, Knoppers BM, Doiron D, Keers JC, Linksted P, Harris JR, Lachance G, Boileau C, Pedersen NL, Hamilton CM, Hveem K, Borugian MJ, Gallagher RP, McLaughlin J, Parker L, Potter JD, Gallacher J, Kaaks R, Liu B, Sprosen T, Vilain A, Atkinson S a, Rengifo A, Morton R, Metspalu A, Wichmann HE, Tremblay M, Chisholm RL, Garcia-Montero A, Hillege H, Litton J-E, Palmer LJ, Perola M, Wolffenbuttel BHR, Peltonen L, Hudson TJ. Quality, quantity and harmony: the DataSHaPER approach to integrating data across bioclinical studies. Int. J. Epidemiol. 2010 Oct;39 (5):1383–93. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2972444&tool=pmcentrez&rendertype=abstract. Accessed 2014-9-3.

22. Cattelani L, Palumbo P, Palmerini L, Chesani F, Chiari L. FRAT-up Web Application. 2014; http://ffrat.farseeingresearch.eu/. .

23. Oliver D, Papaioannou A, Giangregorio L, Thabane L, Reizgys K, Foster G. A systematic review and meta-analysis of studies using the STRATIFY tool for prediction of falls in hospital patients: how well does it work? Age Ageing 2008 Nov;37 (6):621–7. http://ageing.oxfordjournals.org/cgi/content/abstract/37/6/621. Accessed 2012-4-2.

24. Gates S, Smith LA, Fisher JD, Lamb SE. Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. J. Rehabil. Res. Dev. 2008 Jan;45 (8):1105–16. http://www.ncbi.nlm.nih.gov/pubmed/19235113. Accessed 2012-11-3.

25. Schoene D, Wu SM-S, Mikolaizak AS, Menant JC, Smith ST, Delbaere K, Lord SR. 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 Feb;61 (2):202–8. http://www.ncbi.nlm.nih.gov/pubmed/23350947. Accessed 2013-5-23.

26. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. 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 Jan;14 (1):14. http://www.biomedcentral.com/1471-2318/14/14. Accessed 2014-2-21.

Chapter 4

In this Chapter we present a model obtained after extensive search over the InCHIANTI dataset (2313 samples relative to 976 subjects, 1010 variables). The model issues probabilistic predictions on the number of future falls. We benchmark it against other fall risk scores: history of falls, gait speed, Short Physical Performance Battery, and FRAT-up. We study the tradeoff between two competing requirements for a prognostic tool: to be accurate in its predictions, and to be easy and quick to administer.

Introduction

As highlighted in Chapter 1, the most widely accepted paradigm for fall prevention in community-dwelling older adults consists of three sequential stages: screening for high fall risk; assessment of those at high risk on multiple risk factors; implementation of a tailored intervention [1]. The screening serves to focus time and financial resources only on subjects at increased risk, and to relieve low-risk patients from unnecessary investigations. It is required to be short and easy to administer. The multifactorial assessment is intended to identify the risk factors to be targeted by the intervention. It may be performed by the same healthcare professional administering the screening tool, or by a specialist geriatrician.

Whether most screening algorithms suggested in clinical guidelines are based on a combination of some simple questions about history of falls in the previous 12 months and difficulties in walking or balance (e.g. AGS/BGS and NICE guidelines [2]), other tools are based on information on multiple risk factors for falls (see e.g. the '*Stay Independent*' brochure issued by the CDC [3; 4]). A previous version of the AGS/BGS guideline was testes and it was found to be suboptimal [5; 6]. To the best of our knowledge, no published article reports about the predictive accuracy of the actual versions of these screening algorithms.

This Chapter presents the training and test of a statistical model over the dataset InCHIANTI (2313 samples, 1010 variables). The scope is to obtain an accurate prediction on the number of falls that a subject will experience during a time span of one year after the assessment. We test the model on future falls and we benchmark it against other fall risk indicators: history of falls (expressed as number of falls experienced during the 12 months before the assessment), gait speed, the Short Physical Performance Battery (SPPB) summary score [7; 8], and FRAT-up [9] (presented in Chapter 3).

A multifactorial assessment provides more information than a simple screening tool. This information can improve the accuracy of the prediction but is given at the expense of increased time and financial burden for the patient and the health system. We take the number of variables included in the model as a simple surrogate of the burden of the assessment and study this accuracy-parsimony tradeoff.

Methods

Data

The dataset comes from the InCHIANTI study, that was already presented in part in Chapter 3. InCHIANTI is a population-based cohort study about mobility in the elderly. It consists of four waves, separated by about three years one each other. At each wave the subjects are assessed on a number of different domains and are asked about falls experienced in the previous 12 months. More details about the study design and its rationale can be found in [10; 11] and in Appendix 4. We have extracted samples joining information about assessment variables from each wave, and falls reported at the subsequent wave (future falls). We excluded samples relative to subjects younger than 65 at the time of the assessment, and samples without information about future falls. We thus obtained 2313 samples relative to 976 subjects.

Every variable from a table of interest was manually annotated for its type, either continuous or categorical. We excluded: variables that could be easily detected as non-interesting (e.g. date of the home interview), categorical variables with more than two levels, variables with the percentage of missing values greater than 50%, and variables where the missing-value-imputation procedure did not converge. We added variables computed from other variables and considered of interest ('derived' variables, e.g. total score on the SAFE scale for fear of falling computed from answers to the questionnaire). Table 1 gives an overview of the dataset content and of this variables selection and addition procedure.

Table 1. Number of variables divided by area before and after the variables selection and addition procedure

		Number of variables	
Area	Brief description	Before	After
Home interview	MMSE, ADL, IADL, social network, CESD, sleep habits, pain, incontinence, physical activity, falls, fear of falling, shoes, smoking habits	985	165
Clinical visit and disease adjudication	Familiar and personal clinical history, diagnosed medical conditions	1026	151
Physical exam	FICSIT, Purdue Pegboard, Stairs, Repeated Chair Stands, several walking tests, joints range of motion, lower limb muscle power, muscle strength, SPPB summary performance score	597	380
Instrumental exams	ECG, ENG, anthropometric measures, Eco-Color-Doppler, blood pressure, peripheral quantitative computed tomography, bioelectrical impedance analysis	210	111
Laboratory exams	Blood and urine assays	356	103
Medications	Drug classes	88	82
Derived variables	Variables computed from other variables (e.g. living alone derived from questions about social network; pain obtained from questions about pain in specific body parts, etc.)	0	18

Model development

We fit a model that expresses its prediction as a negative binomial distributions on the number of falls, with mean determined by the output a Poisson Lasso regression [12] and a dispersion coefficient calculated as explained below. Fitting and evaluation are performed with 10-fold cross-validation. The splits of samples in folds are done so that all the samples relative to one same subjects are consistently assigned to the same fold. The 9 folds that are in turn assigned to training are used to fit the Poisson Lasso regression [13] and to calculate the dispersion coefficient. The dispersion coefficient is calculated from the predictions *Yh* issued by the regression model and the observed number of falls *Y*, using maximum likelihood and assuming the number of falls as drawn from a negative binomial distribution with mean equal to *Yhl* (R function *theta.ml* from package *MASS* [14]). The regression model and the dispersion coefficient are used to issue the probabilistic predictions on the test samples.

Missing data are imputed on each fold with a different imputation model [15], using age, sex, and time for walking 7 m at self-selected pace as predictors of missing values. These three variables are associated with a number of indicators of health status. The choice of a restricted number of variables for imputing missing data was driven by simplicity.

History of falls (expressed as the number of falls experienced during the 12 months before the assessment), self-selected gait speed as measured in a 7m walk test, and SPPB summary score are all variables already present in the dataset. FRAT-up risk score is calculated as done in [9] and presented in Chapter 3.

Model assessment

We label as fallers (respectively, multiple fallers) all the samples that report at least one fall (two falls) at the follow-up after the assessment. We calculate receiver operating characteristic (ROC) curves of the Lasso model and the other risk scores for fallers and multiple fallers. The discriminative ability is measured with the area under the ROC curve (AUC). The AUC's are compared with bootstrap tests for paired ROC curves [16]. The choice of AUC for comparing the risk scores has a twofold motivation. First of all, it can be employed on non-probabilistic risk scores, as long as they are equipped with an order relation. Thus, it is suitable for gait speed, history of falls, and the SPPB score. Secondly, AUC is commonly used and well known in the field.

The Lasso model is also evaluated for its calibration (i.e. the agreement between its predictions and the observed number of falls) by means of reliability diagram, marginal calibration plot, and probability integral transform (PIT). The reliability diagram (also known as calibration plot or attribute diagram) [17] is adapted for count data, plotting the observed fall rate against the predicted fall rate in samples grouped according to deciles of the expected number of falls. Marginal calibration is evaluated following [18], and defining the relative error for each possible outcome (number of falls from 0 to 9) as the difference between observed and predicted number of samples with that outcome, divided by the total number of samples. The probability integral transform is calculated according to the non-randomized procedure for count data described in [18].

Accuracy-parsimony tradeoff

The Lasso regression performs at the same time variable selection and parameter estimation. It encourages sparse solutions [12; 19]. In order to study how the parsimony of the model impacts on its predictive accuracy, we repeat the whole cross-validation procedure multiple times, under constraints on the total number of variables to include in the model. We evaluate the accuracy of the models obtained under the different constraints. Besides AUC's for single and multiple fallers, mean square error is employed as

additional measure of performance, because it is easily interpretable and is a scoring function that is consistent for the mean, the mean number of falls being in turn the functional elicited during regression [20].

Results

The 10 Lasso models fitted within the procedure of 10-fold cross-validation make use of a number of variables that ranges between 21 and 41, its mean being 29.4. Details on which variables have been selected more frequently and their regression coefficients are given in Appendix 4.

ROC curves of the tools for single and multiple fallers are shown in Figure 2. The associated AUC's and the results of the hypothesis tests for paired ROC curves are shown in Table 2.

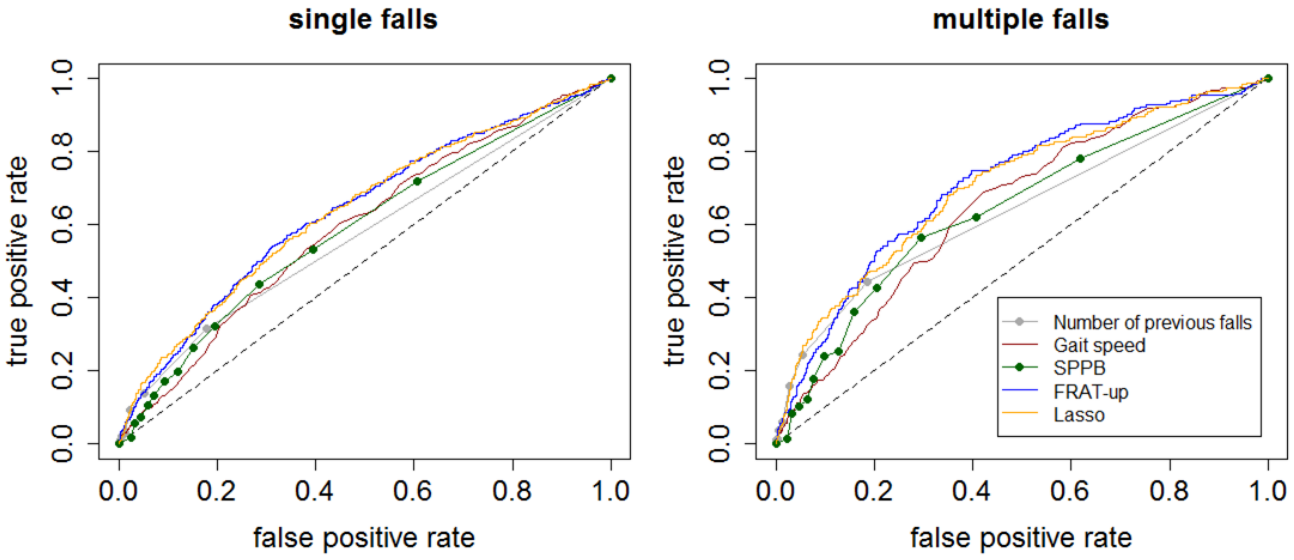

Figure 1. ROC curves for single falls (left) and multiple falls (right).

Table 2. Discriminative ability of five scores for fall risk. Comparison with FRAT-up and Lasso model are drawn with DeLong tests for paired AUC's. * = p<0.01, ** = p<0.001.

	single falls		multiple falls	
	AUC (95% C.I.)	p value vs FRAT-up / Lasso	AUC (95% C.I.)	p value vs FRAT-up/Lasso
Number of previous falls	0.574 (0.551-0.597)	** / **	0.640 (0.603-0.678)	** / **
Gait speed	0.594 (0.566-0.622)	** / **	0.653 (0.615-0.692)	** / *
SPPB	0.590 (0.563-0.618)	** / **	0.645 (0.604-0.686)	** / **
FRAT-up	0.638 (0.610-0.666)	– / 0.92	0.713 (0.675-0.752)	– / 0.62
Lasso	0.639 (0.611-0.667)	0.92 / –	0.708 (0.669-0.747)	0.62 / –

Figure 2 shows an example of the output of the Lasso model for four representative samples at the 2.5th, 10th, 90th, and 97.5th percentiles of the Lasso risk score, compared with the distribution of the number of falls in the InCHIANTI dataset.

Reliability diagram, marginal calibration plot, and histogram of the PIT for the Lasso model are given in Figure 3. Results for the assessment of marginal calibration are shown more in detail in Appendix 4.

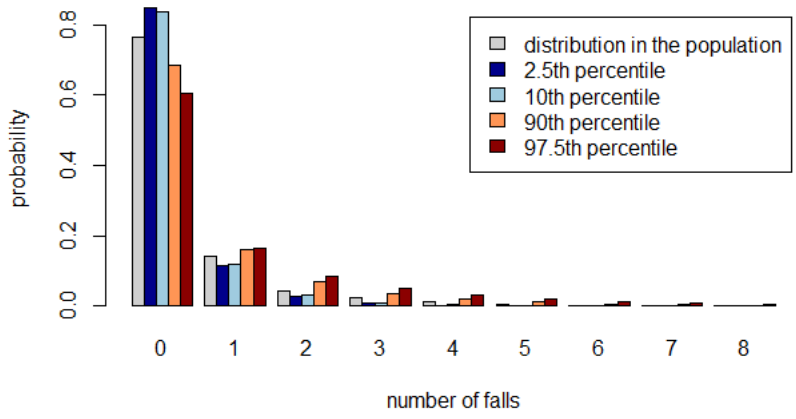

Figure 2. Histogram showing the predictive distributions (probability mass functions) on the number of falls for four samples, being at the 2.5th, 10th, 90th, and 97.5th percentiles of the Lasso risk score. These are compared with the distribution of the number of falls in the InCHIANTI population. The expected number of falls for the four selected samples is respectively 0.21, 0.23, 0.66, 1.08. The fall rate in the InCHIANTI population (baseline data) is 0.42 falls/(person · year).

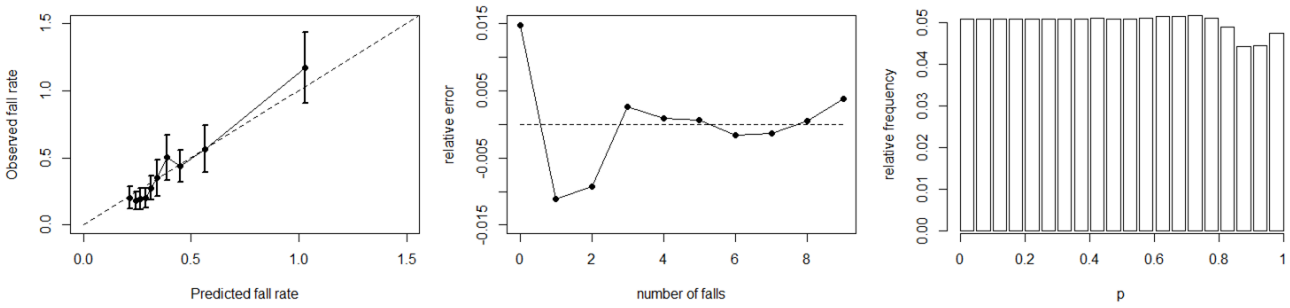

Figure 3. Plots assessing the calibration of the Lasso model. Left: reliability diagram; observed vs predicted fall rate, obtained from grouping samples according to deciles on the risk score; error bars indicate 95% confidence intervals. Center: marginal calibration plot; relative error vs number of falls. Right: histogram of the probability integral transform.

Figure 4 reports the results of the accuracy-parsimony tradeoff analysis. The mean number of variables actually included in the regression model (mean over the 10 models obtained in 10-fold cross-validation) increases and the predictive accuracy (measured with AUC for single and multiple fallers, and mean squared error) improves when relaxing the constraint on the maximum number of variables to include. The AUC's reach a plateau at about 20 features, the MSE at about 30 features.

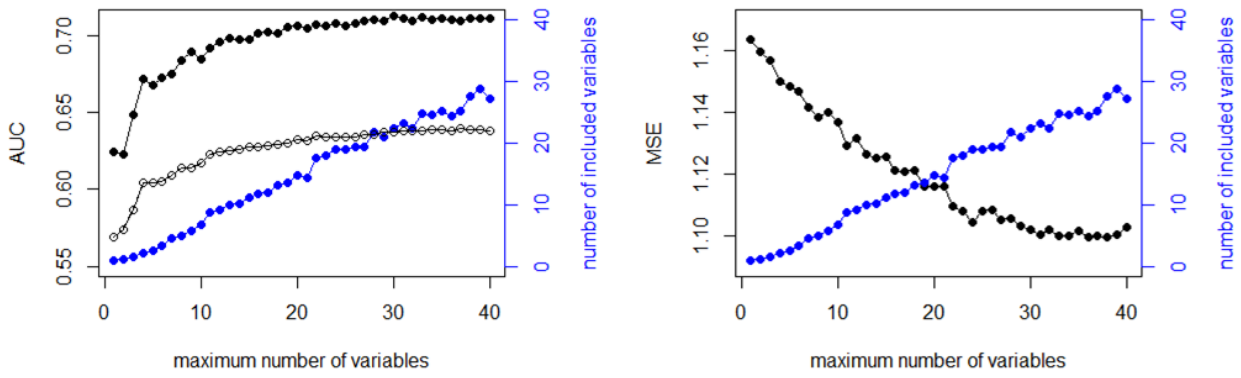

Figure 4. Performance of the model when constraining the maximum number of variables to be possibly included in the model. Left: AUC for single falls (black empty circles), AUC for multiple falls (black filled circles). Right: Mean squared error (MSE, black filled circles). The mean number of variables that were actually included in the models (left and right panels, blue circles) refer to the axes on the right.

Discussion

In this Chapter we have developed a model for fall prediction in a dataset that is large both in terms of samples and of variables related to mobility. We have assessed its predictive properties and have benchmarked it against four other risk scores. We have further investigated whether and to which degree the parsimony of the model, that is required to have a short fall risk screening test, compromises its predictive accuracy.

The four risk scores used for benchmark were: history of falls (expressed as number of falls experienced during the 12 months before the assessment), self-selected gait speed measured in a 7 m walking test, the SPPB summary score[7], and FRAT-up. History of falls is a strong risk indicator for future falls [1; 21; 22]. Clearly, if considered alone, it cannot be used for primary prevention. Gait speed is an indicator of health state in geriatric populations [23]. Its prognostic value for future falls has been shown to be equivalent to total time to perform the TUG [24]. SPPB is a mobility functional test commonly included in comprehensive geriatric assessments [7; 8]. Its association with falls and injurious falls is documented [25; 26]. FRAT-up is the tool proposed in Chapter 3 [9].

The results show that the AUCs of the Lasso model and of FRAT-up are similar and significantly higher than the other risk scores. FRAT-up parameters were derived from the literature, while the approach here proposed is strongly data-driven. The equivalence of discriminative ability between Lasso and FRAT-up validates the literature-driven approach. More analyses presented in Appendix 4 show that the possible dangers of training statistical models in high-dimensional spaces have been avoided. The learning curves, in particular, show that it is not likely to achieve substantial improvements on the AUC having the availability of more samples for training.

The Lasso model results to be well calibrated (Figure 3). Calibration refers to different properties of statistical consistency between predictions and observations [27; 28]. The reliability diagram shows that the number of predicted falls agrees with the number of observed falls on samples grouped in deciles of the risk score. It also shows that discrimination among those at low risk is rather poor. The marginal calibration plot shows that the model performs fairly well in predicting how many subjects will experience a given number of falls (see also Table 3 of Appendix 4). The histogram of the PIT shows that the model is neutrally

dispersed, i.e. the negative binomial distributions meant to express the predictions have a variance that reflects the right amount of uncertainty on the number of falls that the subject will experience [18]. Poisson predictions, obtained without calculating the dispersion coefficient, would substantially underestimate the number of non-fallers and exhibit under-dispersion (results not shown). Given the unexplained variance in falls incidence across different studies [29], the good calibration properties obtained on the InCHIANTI are not guaranteed to hold on other datasets.

Gait speed and the SPPB score (and history of falls) can be interpreted as positively (negatively)-oriented performance scores, so that the lower (respectively, the higher), the higher the risk of falling. FRAT-up and the Lasso model are instead probabilistic predictions. FRAT-up outputs the probability of falling at least once during a time-span of 12 months after the assessment. The Lasso model supplies the distribution on the number of falls that will be experienced during the same time-span. Predicting the number of falls instead of a dichotomous outcome, as the occurrence of at least one or two falls, gives more information without drawbacks.

Expressing a prediction in probabilistic terms has advantages. Firstly, probabilistic predictions allow comparing and aggregating the outcomes of different tools relatives to the same health outcome, and comparing the risks of different health outcomes. Secondly, calibrated probabilistic models allow to make accurate statements on groups of subjects. Furthermore, probabilistic tools do not need any tool-specific knowledge by the user about the semantics of the output. From a psychological perspective, research is currently investigating about the best ways to convey predictions, their associated uncertainty, and their most influential determinants [30–34]. Usable graphic aids could make a difference in the widespread uptake of these prognostic tools.

The accuracy-parsimony analysis shows that screening tools employing a very small number of variables have suboptimal performance. About 20-30 items are required to have good discrimination and accuracy. The AGS/BGS guideline suggests a cascade of 1-4 questions and simple assessments: two or more falls in previous 12 months, acute fall, self-reported difficulties and assessed abnormalities in gait and balance. The '*Stay independent*' brochure [3] by the CDC is made of 12 questions. Its score is integrated with three questions directly asked by the clinician and possibly an assessment of gait, balance and strength [4]. FRAT-up questionnaire is made of 28 items, with the possibility to leave some fields blank [35]. Cognitively functional patients can start answering the questionnaire autonomously, while in the waiting room or even at home through a web application. As this information is generally already collected during a geriatric comprehensive assessment, a fall risk evaluation integrated within this framework should not represent an additional burden to patients and healthcare professionals. Therefore, the choice to feed the prognostic tool with more than 1-4 variables does not conflict with the need of making rational use of time in clinical practice.

The InCHIANTI dataset allowed us to make an extensive search on different domains related to mobility and falls in the elderly. However, we have to acknowledge two main limitations. Firstly, we did not have information about environmental hazards. Secondly, samples were done collating information about risk factors from one wave of the study, and information about future falls from the subsequent wave. As each wave follows the preceding one of about three years, and since at each wave subjects are asked about falls occurred within the last 12 months, predictions here have been made on falls occurring at about 24-36 months after the risk factor assessment. This may be the cause of the relatively low values achieved on the AUC.

Conclusions

We have presented the development and assessment of a tool that issues probabilistic predictions on the number of future falls. We have trained this model over a dataset, that is large both in terms of number of samples and number of variables related to mobility and falls. We have benchmarked it against other risk scores. After extensive search and using state-of-the-art tools of statistical learning, we were not able to reach a better discriminative ability than FRAT-up.

An accuracy-parsimony analysis has highlighted that simplistic screening tests (1-4 variables) are suboptimal in terms of predictive accuracy. Integration of prognostic tools for falls within a geriatric comprehensive assessment can improve the prediction without compromising usability.

References

1. Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ. Will my patient fall? JAMA 2007 Jan;297 (1):77–86. http://www.ncbi.nlm.nih.gov/pubmed/17200478. Accessed 2013-2-1.

2. Panel of Falls in Older Persons American Geriatrics Society and British Geriatrics Society. Summary of the Updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. J. Am. Geriatr. Soc. 2011 Jan;59 (1):148–57. http://www.ncbi.nlm.nih.gov/pubmed/21226685. Accessed 2012-3-14.

3. Centers for Disease Control and Prevention. Stay independent brochure. [date unknown]; http://www.cdc.gov/homeandrecreationalsafety/pdf/steadi/stay_independent.pdf. Accessed 2015-2-21.

4. Centers for Disease Control and Prevention. Algorithm for fall risk assessment & intervention. [date unknown]; http://www.cdc.gov/homeandrecreationalsafety/pdf/steadi/algorithm_fall_risk_assessment.pdf. Accessed 2015-2-21.

5. Lamb SE, McCabe C, Becker C, Fried LP, Guralnik JM. The optimal sequence and selection of screening test items to predict fall risk in older disabled women: the Women's Health and Aging Study. J. Gerontol. Med. Sci. 2008;.

6. Muir SW, Berg K, Chesworth B, Klar N, Speechley M. Application of a fall screening algorithm stratified fall risk but missed preventive opportunities in community-dwelling older adults: A prospective study. J. Geriatr. Phys. Ther. 2010;33 (4):165–172. http://www.scopus.com/inward/record.url?eid=2-s2.0-79955118394&partnerID=tZOtx3y1. .

7. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, Wallace RB. 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;.

8. 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 Mar;332 (9):556–61. http://www.ncbi.nlm.nih.gov/pubmed/7838189. Accessed 2015-1-28.

9. Cattelani L, Palumbo P, Palmerini L, Bandinelli S, Becker C, Chesani F, Chiari L. FRAT-up, a web-based fall risk assessment tool for elderly people living in the community. J. Med. Internet Res. 2015;17 (2).

10. Ferrucci L, Bandinelli S, Benvenuti E, Di Iorio A, Macchi C, Harris TB, Guralnik JM. 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 Dec;48 (12):1618–25. http://www.ncbi.nlm.nih.gov/pubmed/11129752. Accessed 2013-3-11.

11. InCHIANTI. InCHIANTI Study door. [date unknown]; http://inchiantistudy.net/. Accessed 2013-7-6.

12. Tibshirani R. Regression shrinkage and selection via the Lasso. J. R. Stat. Soc. Ser. B (Statistical Methodol. 1996;58 (1):267–288.

13. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. 2010 Jan;33 (1):1–22. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2929880&tool=pmcentrez&rendertype=abstract. Accessed 2015-1-13.

14. Venables VN, Ripley BD. Modern Applied Statistics with S. Fourth. New York: Springer; 2002. http://www.stats.ox.ac.uk/pub/MASS4. .

15. Van Buuren S, Groothuis-oudshoorn K. mice : Multivariate Imputation by Chained Equations in R. J. Stat. Softw. 2011;45 (3).

16. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011 Jan;12 (1):77. http://www.biomedcentral.com/1471-2105/12/77. Accessed 2013-2-10.

17. Murphy AH, Winkler RL. Reliability of subjective probability forecasts of precipitation and temperature. J. R. Soc. Ser. C (Applied Stat. 1977;41–47. http://www.jstor.org/stable/2346866?seq=2#page_scan_tab_contents. Accessed 2015-2-11.

18. Czado C, Gneiting T, Held L. Predictive model assessment for count data. Biometrics 2009 Dec;65 (4):1254–61. http://www.ncbi.nlm.nih.gov/pubmed/19432783. Accessed 2014-3-24.

19. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data Mining, Inference, and Prediction. Second edi. Springer; 2009.

20. Gneiting T. Making and Evaluating Point Forecasts. J. Am. Stat. Assoc. 2011 Jun;106 (494):746–762. http://amstat.tandfonline.com/doi/abs/10.1198/jasa.2011.r10138#.VIWK7TGG8lI. Accessed 2014-12-3.

21. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk Factors for Falls in Community-dwelling Older People: A Systematic Review and Meta-analysis. Epidemiology 2010;21 (5):658–668.

22. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults : A review of the literature. Maturitas 2013;75 (1):51–61. http://dx.doi.org/10.1016/j.maturitas.2013.02.009. .

23. Abellan van Kan G, Rolland Y, Andrieu S, Bauer J, Beauchet O, Bonnefoy M, Cesari M, Donini LM, Gillette Guyonnet S, Inzitari M, Nourhashemi F, Onder G, Ritz P, Salva A, Visser M, Vellas B. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. J. Nutr. Health Aging 2009 Dec;13 (10):881–9. http://www.ncbi.nlm.nih.gov/pubmed/19924348. Accessed 2015-2-5.

24. Viccaro LJ, Perera S, Studenski S a. Is timed up and go better than gait speed in predicting health, function, and falls in older adults? J. Am. Geriatr. Soc. 2011 May;59 (5):887–92.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3522463&tool=pmcentrez&rendertype=abstract. Accessed 2013-5-22.

25. Veronese N, Bolzetta F, Toffanello ED, Zambon S, De Rui M, Perissinotto E, Coin A, Corti M-C, Baggio G, Crepaldi G, Sergi G, Manzato E. Association between Short Physical Performance Battery and falls in older people: the Progetto Veneto Anziani Study. Rejuvenation Res. 2014 Jun;17 (3):276–84. http://www.ncbi.nlm.nih.gov/pubmed/24387140. Accessed 2015-2-18.

26. Ward RE, Leveille SG, Beauchamp MK, Travison T, Alexander N, Jette AM, Bean JF. Functional performance as a predictor of injurious falls in older adults. J. Am. Geriatr. Soc. 2015 Feb;63 (2):315–20. http://www.ncbi.nlm.nih.gov/pubmed/25688606. Accessed 2015-2-19.

27. Gneiting T, Balabdaoui F, Raftery AE. Probabilistic forecasts , calibration and sharpness. J. R. Stat. Soc. Ser. B 2007;69:243–268.

28. Gneiting T, Katzfuss M. Probabilistic Forecasting. Annu. Rev. Stat. Its Appl. 2014;.

29. Rapp K, Freiberger E, Todd C, Klenk J, Becker C, Denkinger M, Scheidt-Nave C, Fuchs J. Fall incidence in Germany: results of two population-based studies, and comparison of retrospective and prospective falls data collection methods. BMC Geriatr. 2014 Jan;14:105. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4179843&tool=pmcentrez&rendertype=abstract. Accessed 2014-11-18.

30. Joslyn S, Leclerc J. Decisions with uncertainty : the glass half full. Curr. Dir. Psychol. Sci. 2013;22 (4):308–315.

31. Spiegelhalter D, Pearson M, Short I. Visualizing uncertainty about the future. Science 2011 Sep;333 (6048):1393–400. http://www.sciencemag.org/content/333/6048/1393.abstract. Accessed 2014-10-20.

32. Edwards A. Explaining risks: turning numerical data into meaningful pictures. BMJ 2002 Apr;324 (7341):827–830. http://www.bmj.com/content/324/7341/827.1?ijkey=685919c5c9016eeb3b049538f47f71263b6aa47e&keytype2=tf_ipsecsha. Accessed 2015-1-26.

33. Smith AF. Discussion of risk pervades doctor-patient communication. BMJ 2002 Sep;325 (7363):548–548. http://www.bmj.com/content/325/7363/548.1. Accessed 2015-2-13.

34. Simila H, Immonen M. Disease state fingerprint for fall risk assessment. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf. 2014 Aug;2014:3176–9. http://www.ncbi.nlm.nih.gov/pubmed/25570665. Accessed 2015-2-13.

35. Cattelani L, Palumbo P, Palmerini L, Chesani F, Chiari L. FRAT-up Web Application. 2014; http://ffrat.farseeingresearch.eu/. .

Chapter 5

This chapter presents an exploratory study on employment of features extracted from wearable inertial sensors signals for fall prediction. Data have been collected within the FARSEEEING project on subject of the InCHIANTI study [1].

Introduction

Wearable inertial sensors have been proposed to increase the accuracy in predicting future falls while guaranteeing the objectivity of the assessment [2]. An overview of the state of the art has already been given in Chapter 1 and Appendix 1. Table 1 below lists all the studies on fall risk that make use of information from wearable inertial sensors and that are based on falls observed prospectively, i.e. during a time-span after the assessment. Other previous studies were based on falls assessed retrospectively [3; 4].

Table 1. Studies using prospective falls

Author	Year	Reference	Number of fallers	Total number of subjects	Notes
Marschollek et al.	2011	[5; 6]	19	46	
Paterson et al.	2011	[7]	54	97	Community-dwelling older women
Schwesig et al.	2013	[8]	17	141	Faller= at least 3 falls in 12 months
Doi et al.	2013	[9]	16	73	
Weiss et al.	2013	[10]	12	71	
Weiss et al.	2014	[11]	14	67	Patients with Parkinson disease
Schwenk et al.	2014	[12]	28	77	Subjects with dementia
Van Schooten et al.	2015	[13]	59	169	

Only few studies have tested the accuracy of their predictive models with techniques for internal validation. To the best of our knowledge, only one study has tested a predictive model on an independent sample of subjects [8]. They obtained an area under the ROC curve (AUC) significantly lower than what obtained in the first study (from 0.791 to 0.587).

Methods

Two hundred fifty-seven subjects aged 65 or more and enrolled in the FARSEEING-InCHIANTI study (FU4) [14] performed four motor functional tests while wearing a smartphone: Timed Up and Go (TUG), Romberg, 5-time repeated chair stand (5RCS), 400 m walk path (400m). A six month follow-up survey (monthly telephone interview) was conducted to assess the occurrence of any falls.
Ninety reliable features were computed from the signals recorded by the inertial sensors embedded in the smartphone. The complete list is available in Table 2.

Table 2. Sensor-based features. StW=sit to walk, TtS=turn to sit, RMS=root mean square, HR=harmonic ratio, AP=antero-posterior, ML=medio-lateral, V=vertical, disp=displacement, accel=acceleration, vel=velocity

TUG	Romberg	5RCS	400m walk (mean across the path segments)
Total duration	RMS ML accel	Total duration	Duration straight path
Duration StW	RMS AP accel	Mean stand duration	Velocity straight path
Duration of turn 180°	RMS V accel	Max accel AP stand	Cadence straight path
Duration Turn TtS	RMS ML gyroscope	Max accel V stand	Standard deviation cadence straight path
Duration TtS	RMS AP gyroscope	RMS accel AP stand	Variation coefficient cadence straight path
RMS AP accel StW	RMS V gyroscope	RMS accel ML stand	Coordination index straight path
RMS ML accel StW	RMS ML disp	RMS accel V stand	Normalized jerk score AP straight path
RMS V accel StW	RMS AP disp	Max angular vel AP stand	Normalized jerk score ML straight path
RMS AP gyroscope StW	RMS V disp	Max angular vel ML stand	Normalized jerk score V straight path
RMS ML gyroscope StW	Sway path V disp	Max angular vel V stand	HR AP straight path
RMS V gyroscope StW	Ellipse area disp	Min angular vel AP stand	HR ML straight path
RMS AP gyroscope TtS	Mean velocity ML displ	Min angular vel ML stand	HR V straight path
RMS ML gyroscope TtS	Mean velocity AP displ	Min angular vel V stand	Step regularity index AP straight path
RMS V gyroscope TtS	Mean velocity V displ	Mean sit duration	Step regularity index ML straight path
Mean velocity 180°	Mean angular velocity ML	Max accel AP sit	Step regularity index V straight path
Mean velocity TtS	Mean angular velocity AP	Max accel V sit	Stride regularity index AP straight path
Peak velocity 180°	Mean angular velocity V	RMS accel AP sit	Stride regularity index ML straight path
Peak velocity TtS	Tremor power % ML	RMS accel ML sit	Stride regularity index V straight path

	Tremor power % AP	RMS accel V sit	Symmetry index AP straight path
	Tremor power % V	Max angular vel AP sit	Symmetry index ML straight path
	RHL ML accel	Max angular vel ML sit	Symmetry index V straight path
	RHL AP accel	Max angular vel V sit	Turn duration
	RHL V accel	Min angular vel AP sit	Mean turn vel
		Min angular vel ML sit	Peak turn vel
		Min angular vel V sit	Normalized jerk score turn

Lasso logistic regression, and linear and quadratic discriminant analysis were applied in order to predict the occurrence of any falls during the follow-up. For the linear and quadratic discriminant analysis, a wrapper feature selection was applied. Five-fold cross validation was used for model assessment.

Results

Twenty-five subjects fell at least once during the 6-month follow-up.

Table 3 shows the features selected by the trained models and the achieved AUC. The features that were more recurrently selected by the trained models were: duration of Sit-to-Walk (TUG), root mean square (RMS) of the acceleration during chair-rising (vertical and anteroposterior components, 5RCS), total time for 5RCS, coefficient of variation of the cadence during straight course (400m).

The AUC's for the trained models range between 0.55 and 0.59, with standard deviations between 0.12 and 0.16.

Table 3. StW=sit to walk, RMS=root mean square, AP=antero-posterior, V=vertical, disp=displacement, accel= acceleration, sd=standard deviation.

	AUC (sd)	Features
Lasso logistic regression	0.58 (0.16)	TUG, Duration StW 5RCS, RMS V accel Romberg, Mean velocity AP displ
Linear discriminant analysis	0.61 (0.12)	5RCS, RMS accel AP stand
Quadratic discriminant analysis	0.61 (0.15)	5RCS, Total duration 400m, Variation coefficient cadence straight path

Discussion and conclusions

The study presented in this chapter is one of the firsts on fall prediction using information on falls collected prospectively. We have trained and internally validated three classifiers, and have identified new possible inertial sensor-based markers of risk of falling.

In line with the low fall rate of the InCHIANTI population, out of 257 subjects, only 25 reported at least one fall at the 6-month follow-up. The low number of fallers could have limited the discriminative accuracy achieved by the classifiers. The AUC may possibly result higher if the follow-up is extended to one year, not only thanks to an increase in cases available for training, but also because of what explained in Chapter 2 [15].

In future studies we will study how to integrate sensor-based information with clinical and behavioral information in order to improve the discriminative ability of the presented models.

All the studies so far have made little use of previous literature. Consequently, despite numerous efforts, we cannot state that the ability to predict falls from wearable sensors has increased during the last years. In our view, the impediments mainly come from the difficulty to align data and results from studies with different designs. We join the plea for more collaborative research [2].

References

1. FARSEEING project; http://farseeingresearch.eu/. .

2. Marschollek M, Schulze M, Gietzelt M, Lovel N, Redmond SJ. Fall prediction with wearable sensors--an empirical study on expert opinions. Stud. Health Technol. Inform. 2013 Jan;190:138–40. http://www.ncbi.nlm.nih.gov/pubmed/23823402. Accessed 2015-2-12.

3. Palumbo P, Palmerini L, Mellone S, Cappello A, Chiari L. Towards a standardized protocol for fall risk assessment using wearable sensors. In: Proceedings of the Third Conference of the Italian National Bioengineering Group (GNB2012). 2012 p. 270–271.

4. Howcroft J, Kofman J, Lemaire ED. Review of fall risk assessment in geriatric populations using inertial sensors. J. Neuroeng. Rehabil. 2013 Aug;10 (1):91. http://www.ncbi.nlm.nih.gov/pubmed/23927446. Accessed 2013-8-16.

5. Marschollek M, Rehwald a, Wolf KH, Gietzelt M, Nemitz G, Meyer Zu Schwabedissen H, Haux R. Sensor-based fall risk assessment--an expert “to go”. Methods Inf. Med. 2011 Jan;50 (5):420–6. http://www.ncbi.nlm.nih.gov/pubmed/21206963. Accessed 2012-11-27.

6. Marschollek M, Rehwald A, Wolf K-H, Gietzelt M, Nemitz G, zu Schwabedissen HM, Schulze M. Sensors vs. experts - a performance comparison of sensor-based fall risk assessment vs. conventional assessment in a sample of geriatric patients. BMC Med. Inform. Decis. Mak. 2011 Jan;11 (1):48. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3141375&tool=pmcentrez&rendertype=abstract. Accessed 2012-5-23.

7. Paterson K, Hill K, Lythgo N. Stride dynamics, gait variability and prospective falls risk in active community dwelling older women. Gait Posture 2011 Feb;33 (2):251–5. http://www.sciencedirect.com/science/article/pii/S0966636210004029. Accessed 2014-12-5.

8. Schwesig R, Fischer D, Lauenroth A, Becker S, Leuchte S. Can falls be predicted with gait analytical and posturographic measurement systems? A prospective follow-up study in a nursing home population. Clin. Rehabil. 2013 Feb;27 (2):183–90. http://www.ncbi.nlm.nih.gov/pubmed/22843355. Accessed 2013-11-8.

9. Doi T, Hirata S, Ono R, Tsutsumimoto K, Misu S, Ando H. The harmonic ratio of trunk acceleration predicts falling among older people: results of a 1-year prospective study. J. Neuroeng. Rehabil. 2013 Jan;10 (1):7. http://www.jneuroengrehab.com/content/10/1/7. Accessed 2014-1-2.

10. Weiss A, Brozgol M, Dorfman M, Herman T, Shema S, Giladi N, Hausdorff JM. Does the Evaluation of Gait Quality During Daily Life Provide Insight Into Fall Risk?: A Novel Approach Using 3-Day Accelerometer Recordings. Neurorehabil. Neural Repair 2013 Jun; http://www.ncbi.nlm.nih.gov/pubmed/23774124. Accessed 2013-6-21.

11. Weiss A, Herman T, Giladi N, Hausdorff JM. Objective assessment of fall risk in Parkinson's disease using a body-fixed sensor worn for 3 days. PLoS One 2014 Jan;9 (5):e96675. http://dx.plos.org/10.1371/journal.pone.0096675. Accessed 2014-11-21.

12. Schwenk M, Hauer K, Zieschang T, Englert S, Mohler J, Najafi B. Sensor-Derived Physical Activity Parameters Can Predict Future Falls in People with Dementia. Gerontology 2014 Aug;60 (6):483–492. http://www.karger.com/Article/FullText/363136. Accessed 2014-9-10.

13. Van Schooten KS, Pijnappels M, Rispens SM, Elders PJM, Lips P, van Dieën JH. Ambulatory Fall-Risk Assessment: Amount and Quality of Daily-Life Gait Predict Falls in Older Adults. J. Gerontol. A. Biol. Sci. Med. Sci. 2015 Jan;glu225–. http://biomedgerontology.oxfordjournals.org/content/early/2015/01/07/gerona.glu225.short. Accessed 2015-1-14.

14. InCHIANTI. InCHIANTI Study door. [date unknown]; http://inchiantistudy.net/. Accessed 2013-7-6.

15. Palumbo P, Palmerini L, Chiari L. A probabilistic model to investigate the properties of prognostic tools for falls. Methods Inf. Med. 2014;53 (6).

Chapter 6

Final remarks

In Chapter 1 we have introduced the topic of this thesis within the wider framework of the demographic trend towards an ageing population and of medical research on prognostic tools. We have further reviewed current clinical guidelines and the scientific literature about fall risk assessment in community-dwelling older adults. The state of the art is characterized by abundancy of newly-developed prognostic tools, paucity of validation studies, and absence of impact analyses.

In Chapter 2 we have developed a probabilistic model that clarifies some conceptual issues about the performance indices of prognostic tools for falls. We have adopted as a model an ideal tool, that is perfect in a probabilistic way, and that issues predictions over subjects of a population on which fall risk and fall counts are distributed according to the classical scheme by Greenwood and Yule for accident proneness. We have shown that the tool does not achieve perfect discrimination. We have also highlighted the sensitivity of some performance indices to key elements of the design and the analysis of validation studies.

In Chapter 3 we have presented FRAT-up, a prognostic tool for falls developed within the framework of the EU project FARSEEING. The core of the tool is represented by the noisy-or gate. Parameter estimation is based exclusively on data from the literature. We have validated FRAT-up on three wide epidemiological studies about ageing: Invecchiare nel Chianti (InCHIANTI), the ActiFE-Ulm study, and the English Longitudinal Study of Ageing (ELSA). Results are heterogeneous among datasets. For example, the area under the ROC curve for single fallers ranges from 0.57 in the ActiFE-Ulm study to 0.70 in ELSA. We have shown that these differences can be found also for other prognostic tools.

In Chapter 4 we have presented a comparison in the same dataset (InCHIANTI) of different fall risk scores: history of falls, gait speed, the Short Physical Performance Battery, FRAT-up, and a prediction tool based on a Lasso regression. FRAT-up and the Lasso model have equivalent discriminative abilities, and they outperform the other simple risk scores. An accuracy-parsimony tradeoff analysis shows that, although short screening tests are preferable because of their ease of administration, disregarding information about risk factors sensibly affects the predictive accuracy.

In Chapter 5 we have presented some results about fall prediction from wearable inertial sensors. Data have been taken from the last follow-up of the FARSEEING-InCHIANTI study.

In addition to what already discussed in the single chapters, we highlight here some general considerations.

Firstly, impact studies quantifying the potential benefits of fall risk assessment are needed. These may be either in the form of randomized controlled trials or may be based on a modelling approach.

Secondly, on the one hand predictive models trained within one same datasets often result to be equivalent in terms of their discriminative ability. This has been shown especially in Chapter 5. Although FRAT-up has not been trained, also FRAT-up and the Lasso model attain similar AUC on the InCHIANTI dataset (Chapter 4). On the other hand, between-dataset differences are remarkable. The cause can be true differences among populations and differences in data collection methods. These differences often are largely unexplained. They hinder the validation of predictive models on external datasets and are responsible for heterogeneity in validation results.

Thirdly, prediction of falls from extremely simple (1-4 questions) algorithms is suboptimal. Falls have a multifactorial etiology. Knowledge about exposure to several risk factors improves the prediction. As this information is generally already collected during a geriatric comprehensive assessment, a fall risk evaluation integrated within this framework should not represent an additional burden to patients and the healthcare professionals. Therefore, the choice to feed the prognostic tool with more than 1-4 variables does not conflict with the need of making rational use of time in clinical practice.

Lastly, research on fall prediction through features extracted from wearable inertial sensors is characterized by a high-dimensional search space and low interpretability of the candidate predictors. This situation requires large datasets in terms of number of samples. If impact studies support the strategic role of fall prediction for fall prevention, efforts should be made to create larger datasets, and to work on standardization and comparability among studies, enabling more collaborative research.

Acknowledgements

This work has been partially supported by Sovvenzione Globale Spinner 2013 (Regione Emilia Romagna) and the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement FARSEEING n° 288940.

The ActiFE-study was performed at Ulm University in cooperation with the Agaplesion Bethesda Clinic Ulm and the study and data center is at the Institute of Epidemiology and Medical Biometry, Ulm University, Ulm, Germany. The study was funded by a grant from the Ministry of Science, Research and Arts, state of Baden-Wuerttemberg, Germany, as part of the Geriatric Competence Center, Ulm University and by funds from the Department of Internal Medicine II-Cardiology, Ulm University.

The data relative to the English Longitudinal Study of Ageing (ELSA) were made available through the UK Data Archive (UKDA). 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 UK government departments co-ordinated 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 (Invecchiare in Chianti, aging in the Chianti area) is currently supported by a grant from the National Institute on Aging (NIH, NIA, Bethesda, USA) 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 two small towns in the countryside of the Tuscany were the study is conducted. The Study was initially managed by the National Institute on Research and Care of the Elderly (INRCA, Ancona, Italy) and it was funded by Italian Health Ministry and by a NIH contract.

I give thanks to my supervisor, all the members of my research group, the group of Artificial Intelligence of the University of Bologna, the group of the InCHIANTI study, and the groups at Robert Bosch Hospital and at the University of Ulm that have given me hospitality during my visit in Germany. All your help has been much appreciated.

Il mio grazie più grande va a mio padre, mia madre, mio fratello e mia sorella.

Appendix 1

This appendix contains details and references about Figure 1 and Figure 7 of Chapter 1.

Figure 1, Chapter 1

Medical outcome	Reference	Tool	Disease	Time span	AUC	M/W
CVD	[1]	Framingham Risk Score	CHD (angina, MI, sudden death)	10 years	0.7744	W
		Framingham Risk Score	CHD (angina, MI, sudden death)	10 years	0.7598	M
		Global cardiovascular risk	CHD, stroke, CHF, PVD	10 years	0.793	W
		Global cardiovascular risk	CHD, stroke, CHF, PVD	10 years	0.763	M
		Adult Treatment Panel III				
		SCORE (Systematic Coronary Risk Evaluation)	fatal CV events (heart attack, stroke, aortic aneurism)	10 years	0.71-0.80	
		Reynolds Risk Score	MI, stroke, coronary revascularization, CV death	10 years	0.808	W
		Reynolds, men	MI, stroke, coronary revascularization, CV death	10 years	0.7-0.714	M
		ASSIGN (Assessing Cardiovascular Risk to Scottish intercollegiate guidelines Network/SIGN to assign preventive treatment)	CV death, CHD admission, CABG, PTCA	10 years	0.7841	W
		ASSIGN (Assessing Cardiovascular Risk to Scottish intercollegiate guidelines Network/SIGN to assign preventive treatment)	CV death, CHD admission, CABG, PTCA	10 years	0.7644	M
		QRISK (QRESEARCH Cardiovascular Risk Algorithm)	MI, stroke, CHD, TIA	10 years	0.7879	W
		QRISK (QRESEARCH Cardiovascular Risk Algorithm)	MI, stroke, CHD, TIA	10 years	0.7674	M
Diabetes	[2]	ARIC (Atherosclerosis Risk in Communities)	type 2 diabetes		0.84	

		Ausdrisk	type 2 diabetes		0.78	
		Cambridge risk score	type 2 diabetes		0.72	
		FINDRISK	type 2 diabetes		0.76	
		Framingham Offspring Study	type 2 diabetes		0.78	
		San Antonio risk score	type 2 diabetes		0.83	
		San Antonio risk score	type 2 diabetes		0.78	
		QDScore	type 2 diabetes		0.8	M
		QDScore	type 2 diabetes		0.81	W
Breast cancer	[3]	Gail 1	breast cancer		NA	
		Gail 2	breast cancer		0.63	
		Colditz and Rosner	breast cancer		0.63	
		Tyrer and Cuzick	breast cancer		0.762	
	[4]	Petracci	breast cancer		0.62	W <50 yr
		Petracci	breast cancer		0.57	W >=50 yr
	[5]	Pfeiffer	breast cancer		0.58	W >=50 yr
Colorectal cancer	[6]	Harvard Cancer Risk Index			0.71	M
		Harvard Cancer Risk Index			0.67	W
		Imperiale			0.74	
		Freedman			0.61	
		Ma			0.64	
Hip fracture	[7]	FRAX	hip fracture	10 years	0.746429	
		FRAX without BMD	hip fracture	10 years	0.717571	
		Garvan FRC	hip fracture		0.773333	
		QFractureScores	hip fracture		0.756	
Falls	[8]	Tinetti B			0.62	
	[9]	Tinetti B		1 year	0.559	
	[10]	TUG			0.66	
	[9]	Functional reach		1 year	0.509	

Meta-analysis performed to assign one value of AUC to Timed-Up and Go Test (TUG). References extracted from [10].

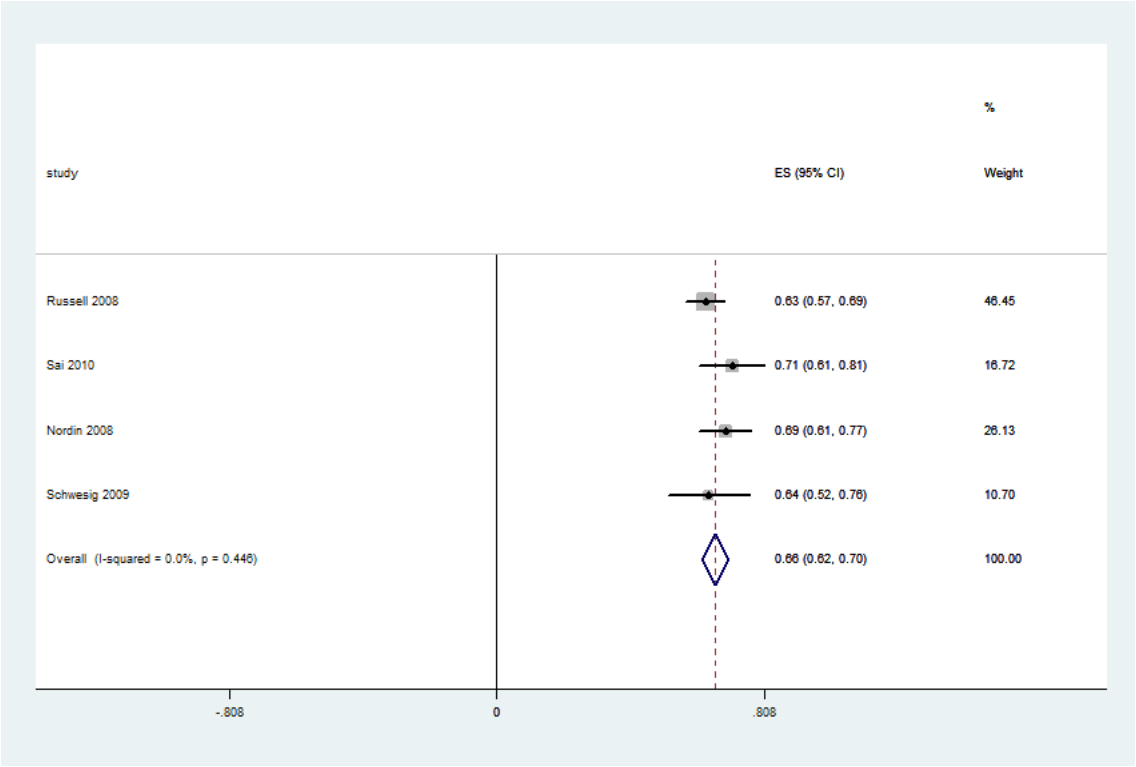

Figure 1. Meta-analysis performed to assign one value of AUC to TUG.

Figure 7, Chapter 1

Traditional tools [8]

Reference	Tool	Sensitivity	Specificity
Bogle Thrbahn	Berg balance test	0.56	0.96
Faber et al 2006	Tinetti total	0.64	0.661
Faber et al 2006	Tinetti Balance	0.64	0.625
Raiche et al 2000	Tinetti total	0.70	0.52
Trueblood 2001	Tinetti balance	0.24	0.91
Trueblood 2001	Tinetti gait	0.21	0.95
Trueblood 2001	TUG	0.10	0.95
Morris et al 2007	5m-TUG	0.949	0.106
		0.718	0.426
		0.513	0.638
		0.385	0.766
		0.333	0.851
		0.205	0.936
		0.128	0.979

Sensor-based tools. Main paper of reference [11]. More than one model was extracted from some paper.

Author	Year	Reference	Number of fallers	Total number of subjects
Caby et al	2011	[12]	15	20
Marschollek et al	2011	[13]	19	46

Marschollek et al	2011	[14]	19	46
O'Sullivan et al	2009	[15]	12	17
Weiss et al	2011	[16]	23	41
Greene et al	2010	[17]	207	349
Riva et al	2013	[18]	42	131
Greene et al	2014	[19]	11	33
			37	91
			48	124

References

1. Berger JS, Jordan CO, Lloyd-Jones D, Blumenthal RS. Screening for cardiovascular risk in asymptomatic patients. J. Am. Coll. Cardiol. 2010 Mar;55 (12):1169–77. http://www.ncbi.nlm.nih.gov/pubmed/20298922. Accessed 2013-10-19.

2. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes : systematic review. Br 2011;7163 (November):1–31.

3. 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 Apr;132 (2):365–77. http://www.ncbi.nlm.nih.gov/pubmed/22037780. Accessed 2013-10-25.

4. Petracci E, Decarli A, Schairer C, Pfeiffer RM, Pee D, Masala G, Palli D, Gail MH. Risk factor modification and projections of absolute breast cancer risk. J. Natl. Cancer Inst. 2011 Jul;103 (13):1037–48. http://jnci.oxfordjournals.org/content/103/13/1037. Accessed 2013-10-25.

5. Pfeiffer RM, Park Y, Kreimer AR, Lacey J V, Pee D, Greenlee RT, Buys SS, Hollenbeck A, Rosner B, Gail MH, Hartge P. Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. PLoS Med. 2013 Jul;10 (7):e1001492. http://dx.plos.org/10.1371/journal.pmed.1001492. Accessed 2013-10-25.

6. Win AK, Macinnis RJ, Hopper JL, Jenkins MA. Risk prediction models for colorectal cancer: a review. Cancer Epidemiol. Biomarkers Prev. 2012 Mar;21 (3):398–410. http://www.ncbi.nlm.nih.gov/pubmed/22169185. Accessed 2013-10-25.

7. Leslie WD, Lix LM. Comparison between various fracture risk assessment tools. Osteoporos. Int. 2014 Jan;25 (1):1–21. http://www.ncbi.nlm.nih.gov/pubmed/23797847. Accessed 2014-4-10.

8. Gates S, Smith LA, Fisher JD, Lamb SE. Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. J. Rehabil. Res. Dev. 2008 Jan;45 (8):1105–16. http://www.ncbi.nlm.nih.gov/pubmed/19235113. Accessed 2012-11-3.

9. Lin M-R, Hwang H-F, Hu M-H, Wu H-DI, Wang Y-W, Huang F-C. 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 Aug;52 (8):1343–8. http://www.ncbi.nlm.nih.gov/pubmed/15271124. Accessed 2013-10-21.

10. Schoene D, Wu SM-S, Mikolaizak AS, Menant JC, Smith ST, Delbaere K, Lord SR. 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 Feb;61 (2):202–8. http://www.ncbi.nlm.nih.gov/pubmed/23350947. Accessed 2013-5-23.

11. Howcroft J, Kofman J, Lemaire ED. Review of fall risk assessment in geriatric populations using inertial sensors. J. Neuroeng. Rehabil. 2013 Aug;10 (1):91. http://www.ncbi.nlm.nih.gov/pubmed/23927446. Accessed 2013-8-16.

12. Caby B, Kieffer S, de Saint Hubert M, Cremer G, Macq B. Feature extraction and selection for objective gait analysis and fall risk assessment by accelerometry. Biomed. Eng. Online 2011 Jan;10 (1):1. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3022766&tool=pmcentrez&rendertype=abstract. Accessed 2012-4-5.

13. Marschollek M, Rehwald A, Wolf K-H, Gietzelt M, Nemitz G, zu Schwabedissen HM, Schulze M. Sensors vs. experts - a performance comparison of sensor-based fall risk assessment vs. conventional assessment in a sample of geriatric patients. BMC Med. Inform. Decis. Mak. 2011 Jan;11 (1):48. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3141375&tool=pmcentrez&rendertype=abstract. Accessed 2012-5-23.

14. Marschollek M, Rehwald a, Wolf KH, Gietzelt M, Nemitz G, Meyer Zu Schwabedissen H, Haux R. Sensor-based fall risk assessment--an expert “to go”. Methods Inf. Med. 2011 Jan;50 (5):420–6. http://www.ncbi.nlm.nih.gov/pubmed/21206963. Accessed 2012-11-27.

15. O’Sullivan M, Blake C, Cunningham C, Boyle G, Finucane C. Correlation of accelerometry with clinical balance tests in older fallers and non-fallers. Age Ageing 2009 May;38 (3):308–13. http://www.ncbi.nlm.nih.gov/pubmed/19252205. Accessed 2012-4-6.

16. Weiss A, Herman T, Plotnik M, Brozgol M, Giladi N, Hausdorff JM. An instrumented timed up and go: the added value of an accelerometer for identifying fall risk in idiopathic fallers. Physiol. Meas. 2011 Dec;32 (12):2003–18. http://www.ncbi.nlm.nih.gov/pubmed/22094550. Accessed 2012-6-18.

17. Greene BR, O’Donovan A, Romero-Ortuno R, Cogan L, Scanaill CN, Kenny R a. Quantitative falls risk assessment using the timed up and go test. IEEE Trans. Biomed. Eng. 2010 Dec;57 (12):2918–26. http://www.ncbi.nlm.nih.gov/pubmed/20923729. .

18. Riva F, Toebes MJP, Pijnappels M, Stagni R, van Dieën JH. Estimating fall risk with inertial sensors using gait stability measures that do not require step detection. Gait Posture 2013 Jun;38 (2):170–4. http://www.sciencedirect.com/science/article/pii/S0966636213002191. Accessed 2014-3-4.

19. Greene BR, Doheny EP, Kenny RA, Caulfield B. Classification of frailty and falls history using a combination of sensor-based mobility assessments. Physiol. Meas. 2014 Oct;35 (10):2053–66. http://stacks.iop.org/0967-3334/35/i=10/a=2053. Accessed 2014-11-7.

Appendix 2

Here we recall the formula for the probability mass function of a negative binomial distribution and how it arises as a mixture of Poisson distributions with a gamma mixing distribution.

The probability mass function of a negative binomial distribution is

$$h(n) = \frac{\Gamma(n+k)}{n!\Gamma(k)}\left(\frac{\mu}{\mu+k}\right)^n\left(\frac{k}{\mu+k}\right)^k, n = 0,1,2,\ldots \qquad \text{(A1)}$$

The mean of this distribution is μ and the variance is $\mu + \frac{\mu^2}{k}$. As the parameter k increases, the variance shrinks toward the mean and the distribution approaches a Poisson distribution. Accordingly, $1/k$ is often referred to as the dispersion parameter.

Given the conditional distribution for N stated in (1), the assignment of a gamma distribution with shape and scale parameters respectively k' and θ for the conditioning variable Λ, and the expression of its probability density function recalled in (5) (with k substituted by k'), the marginal distribution of N is

$$P(N = n) = \int_0^{+\infty} g(n;\lambda\tau)f(\lambda)d\lambda = \int_0^{+\infty} e^{-\lambda\tau}\frac{(\lambda\tau)^n}{n!}\frac{\lambda^{k\prime-1}e^{-\lambda/\theta}}{\Gamma(k')\theta^{k\prime}}d\lambda$$

$$= \frac{\tau^n}{n!\,\Gamma(k')\theta^{k\prime}}\int_0^{+\infty} e^{-\lambda(\tau+1/\theta)}\lambda^{n+k\prime-1}\,d\lambda = \frac{\tau^n}{n!\,\Gamma(k')\theta^{k\prime}}\left(\frac{\theta}{\tau\theta+1}\right)^{n+k\prime}\int_0^{+\infty} e^{-s}s^{n+k\prime-1}\,ds$$

$$= \frac{\Gamma(n+k\prime)}{n!\Gamma(k\prime)}\left(\frac{\theta\tau}{\theta\tau+1}\right)^n\frac{1}{(\theta\tau+1)^{k\prime}} \qquad \text{(A2)}$$

After comparison between (A1) and (A2) we see that marginally N follows a negative binomial distribution, whose parameter k is equal to the shape parameter of the gamma distribution k' and whose parameter μ is related to the shape and scale parameters of the gamma distribution k' and θ, and to the parameter τ by the relation $\mu = \theta\tau k'$.

The identifiability of Poisson mixture distribution allows to state that the gamma distribution is the only mixing distribution that makes the mixture follow a negative binomial distribution. More properties of Poisson mixture distributions are reviewed in [1].

References

1. Karlis D, Xekalaki E. Mixed Poisson Distributions. Int. Stat. Rev. 2005 Jan;73 (1):35–58. http://doi.wiley.com/10.1111/j.1751-5823.2005.tb00250.x. .

Appendix 3

This Appendix contains material supplemental to Chapter 3. In particular, a description of FRAT-up risk factors, analyses on the harmonized datasets, a sensitivity analysis on FRAT-up performance. Details about source variables, target variables, and harmonization algorithms for the three datasets are given in Appendix 3-ActiFE, Appendix 3-ELSA, and Appendix 3-InCHIANTI.

Risk factors

There are two types of risk factors, either dichotomous (supporting missing data, so the possible values are "true", "false", and "unknown") and scalar (integer values with the possibility of being unknown). Scalar values have a range n,…,m, so a scalar 1,…,3 may take the values 1,2,3, and "unknown".

FRAT-up supports the same risk factors that were found to be significant in the Deandrea meta-analysis [1].

Age, number of medications, and comorbidity are scalar risk factors. Age increases of a level every five years, starting from level zero at age 65 (e.g. it becomes level 1 at 70), with 4 as the maximum level for subjects aged 85 or more. Number of medications ranges from 0 to 10. For N=0,…,9, it takes level N on subjects taking N drugs. It takes level 10 on subjects taking 10 or more drugs.

Comorbidity counts the number of morbid conditions from the following list of 11 risk factors: cognition impairment, depression, diabetes, dizziness and vertigo, fear of falling, history of stroke, pain, Parkinson disease, poor self-perceived health status, rheumatic disease, and urinary incontinence. More precisely, it assumes level 0 if there are 0 or 1 exposures, and level N-1 otherwise, with N being the number of morbid conditions. It thus can range from 0 to 10.

Fall risk factor prevalence from literature

In Table 1, the "Prevalence" column contains the probability to be "true" of dichotomous risk factors. The column contains the prevalence of the single levels for scalar risk factors, ordered from level zero upward.

Table 1. Risk factor prevalence from literature, sources and notes

Name	Prevalence	Source	Notes
age	0.25, 0.25, 0.20, 0.16, 0.14	[2]	The distribution of age (divided in five years intervals) in the Italian population
cognition impairment	0.19	[3]	SPMSQ ≥ 3
comorbidity	Inferred		
depression	0.13	[4]	CES-D ≥ 16

diabetes	0.11	[5]	
dizziness and vertigo	0.20	[6]	
fear of falling	0.33	[7]	
female sex	0.48	[6]	
gait problems	0.42	[8]	Difficulty walking
hearing impairment	0.36	[4]	Questioning the participant on whether he/she could follow a conversation in a group of four persons (with a hearing aid if needed)
history of falls	0.31	[9]	People 65+ having fallen at least once in 12 months
history of stroke	0.13	[5]	
instrumental disability	0.37	[5]	One or more IADL impairment
living alone	0.32	[6]	
number of medications	0.237, 0.226, 0.194, 0.133, 0.081, 0.049, 0.036, 0.02, 0.01, 0.007, 0.007	[10]	Using medications in past two days, by number of medications, household population aged 65 or older, Canada excluding territories
Pain	0.30	[11]	
parkinson	0.008	[5]	
physical activity limitation	0.56	[12]	Self-reported physical activity levels in adults, by sex and age, England 2008, low activity: less than 30 minutes or more of moderate or vigorous activity on 1 to 4 days a week
physical disability	0.11	[3]	ADL ≤ 4
poor self-perceived health status	0.20	[6]	Poor subjective health status (≥ 4)
rheumatic disease	0.47	[13]	Arthritis
urinary incontinence	0.19	[6]	
use of antiepileptics	0.01	[4]	
use of	0.32	[7]	

antihypertensives			
use of sedatives	0.14	[4]	Use of benzodiazepines
vision impairment	0.19	[4]	Questioning the participant on whether he/she could recognize someone's face at a distance of 4 meters (with glasses or contact lenses if needed)
walking aid use	0.18	[6]	

Harmonized datasets

Missing values

Table 2 gives the frequencies of missing values on the harmonized variables in the three datasets before and after excluding those subjects on which there is no information on the outcome variable.

Table 2. Frequency of missing values before and after excluding subjects on which there is no information on the outcome variables.

	ActiFE		InCHIANTI		ELSA	
	Before exclusion	After exclusion	Before exclusion	After exclusion	Before exclusion	After exclusion
Age	0.000	0.000	0.000	0.000	0.000	0.000
Sex	0.000	0.000	0.000	0.000	0.000	0.000
History of falls (yes/no)	0.017	0.013	0.000	0.000	0.001	0.001
History of falls (number)	0.056	0.053	0.004	0.002	0.004	0.004
Prospective falls (yes/no)	0.060	0.000	0.225	0.000	0.195	0.000
Prospective falls (number)	0.060	0.000	0.228	0.000	0.199	0.000
Living alone	0.015	0.016	0.000	0.000	0.000	0.000
Walking aid use	0.088	0.085	0.112	0.064	0.012	0.010
Urinary incontinence	0.019	0.014	0.002	0.000	0.001	0.001
Diabetes	0.003	0.003	0.000	0.000	0.000	0.000
Parkinson disease	0.000	0.000	0.010	0.008	0.000	0.000
Arthritis or rheumatism	0.003	0.004	0.000	0.000	0.000	0.000
Cognition impairment	0.082	0.081	0.000	0.000	0.000	0.000
History of stroke	0.003	0.004	0.003	0.002	0.000	0.000
Depression	0.063	0.052	0.066	0.025	0.000	0.000
Poor self-perceived health status	0.005	0.005	0.067	0.026	1.000	1.000

Pain	0.007	0.006	0.078	0.037	0.007	0.005
Physical disability	0.011	0.010	0.000	0.000	0.000	0.000
Instrumental disability	0.019	0.019	0.000	0.000	0.000	0.000
Fear of falling	0.021	0.014	0.003	0.001	0.001	0.001
Dizziness	0.014	0.010	0.127	0.066	0.027	0.019
Vision impairment	0.022	0.017	0.207	0.146	0.000	0.000
Hearing impairment	0.021	0.017	0.104	0.058	0.000	0.000
Number of medications	0.000	0.000	0.000	0.000	0.005	0.005
Use of antihypertensives	0.000	0.000	0.000	0.000	1.000	1.000
Use of sedatives	0.000	0.000	0.225	0.000	1.000	1.000
Use of antiepileptics	0.000	0.000	0.225	0.000	1.000	1.000
Physical activity limitations	0.155	0.143	0.004	0.003	0.001	0.001
Gait problems	0.029	0.030	0.129	0.080	0.104	0.091

Principal components analysis

Principal component analysis (PCA) is here performed to complete the description of the distribution of the harmonized variables furnished in Chapter 4.

Three PCAs were performed separately on the correlation matrices of the three datasets. The results thereof are shown in Figure 1, Figure 2 and Table 3. One further PCA was performed on the correlation matrix of the three datasets stacked together. The results thereof are showed in Figure 3, Figure 4 and Table 4. The PCAs were performed on one of the 11 imputed copies of the dataset (see the section on missing data imputation). In order to adjust for differential nonresponse rates to the nurse visit in the ELSA population, weights were used to calculate the correlation matrices.

Consistently across the datasets, the first PC has greater loadings on age, scores of the SPPB, physical activity limitation, and physical and instrumental disability. Interpreting it as a cline for functional performance, a ceiling effect is evident from the scatterplots in Figure 2 and Figure 4. The second PCs of the PCAs performed on the three separate datasets, and the third PC of the PCA performed on the datasets stacked together have great loadings on sex and grip strength.

The second PC of the PCA performed on the datasets stacked together discriminates the ActiFE and InCHIANTI populations from ELSA. The greatest loadings to this PC are given by pain, vision impairment and number of medications. This may be an indication that harmonization of these variables was not successful on ELSA.

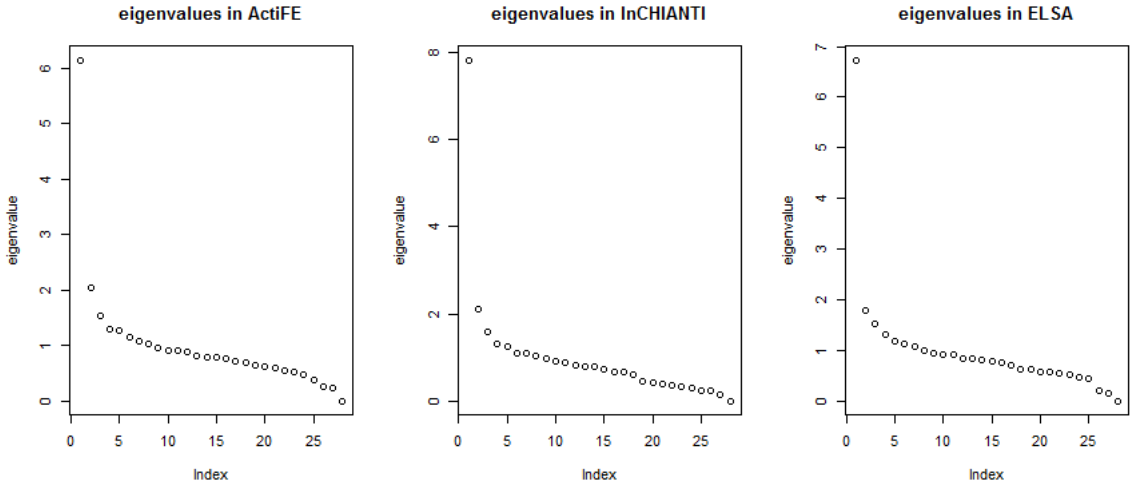

Figure 1. Eigenvalues of the correlation matrices of the three harmonized datasets.

Table 3. Loadings on the harmonized variables of the first and second principal components as calculated from PCAs performed separately on the three datasets.

	Loadings					
	ActiFE		InCHIANTI		ELSA	
	1st PC	2nd PC	1st PC	2nd PC	1st PC	2nd PC
Age	0.205	0.013	0.242	0.031	0.204	-0.195
Sex (women)	0.063	-0.543	0.102	-0.442	0.107	-0.542
Gait speed	-0.287	-0.010	-0.306	-0.029	-0.291	0.035
SPPB_BT score	-0.284	-0.127	-0.289	-0.104	-0.273	0.028
SPPB_GST score	-0.315	-0.100	-0.296	-0.171	-0.269	0.020
SPPB_CST score	-0.290	-0.034	-0.287	-0.069	-0.265	0.005
SPPB score	-0.366	-0.100	-0.328	-0.125	-0.340	0.021
History of falls (yes/no)	0.101	-0.129	0.095	-0.296	0.136	-0.056
History of falls (number)	0.041	0.033	0.107	-0.275	0.074	0.049
Grip strength	-0.160	0.455	-0.221	0.302	-0.216	0.470
Walking aid use	0.245	0.161	0.260	0.209	0.241	0.122
Living alone	0.096	-0.201	0.040	-0.210	0.100	-0.343
Urinary incontinence	0.086	-0.349	0.149	-0.196	0.101	-0.037
Diabetes	0.080	0.077	0.030	0.066	0.067	0.267
Parkinson disease	0.037	0.054	0.046	-0.032	0.046	0.071
Arthritis or rheumatism	0.041	-0.306	0.080	-0.221	0.147	-0.075
Cognition impairment	0.101	0.043	0.195	0.116	0.036	0.067
History of stroke	0.064	0.041	0.095	0.125	0.096	0.142
Depression	0.119	-0.104	0.105	-0.233	0.109	0.017
Pain	0.073	-0.211	0.059	-0.237	0.162	0.143
Physical disability	0.257	0.136	0.234	0.190	0.241	0.125
Instrumental disability	0.263	-0.024	0.271	0.018	0.255	0.118

Fear of falling	0.243	-0.022	0.117	-0.257	0.197	0.097
Dizziness	0.115	-0.220	0.023	-0.248	0.235	0.107
Vision impairment	-0.108	-0.095	0.161	0.032	0.141	0.070
Hearing impairment	0.137	0.071	0.054	0.024	0.084	0.226
Number of medication	0.174	-0.034	0.144	-0.054	0.132	0.229
Physical activity limitations	0.222	0.100	0.231	0.018	0.206	0.141

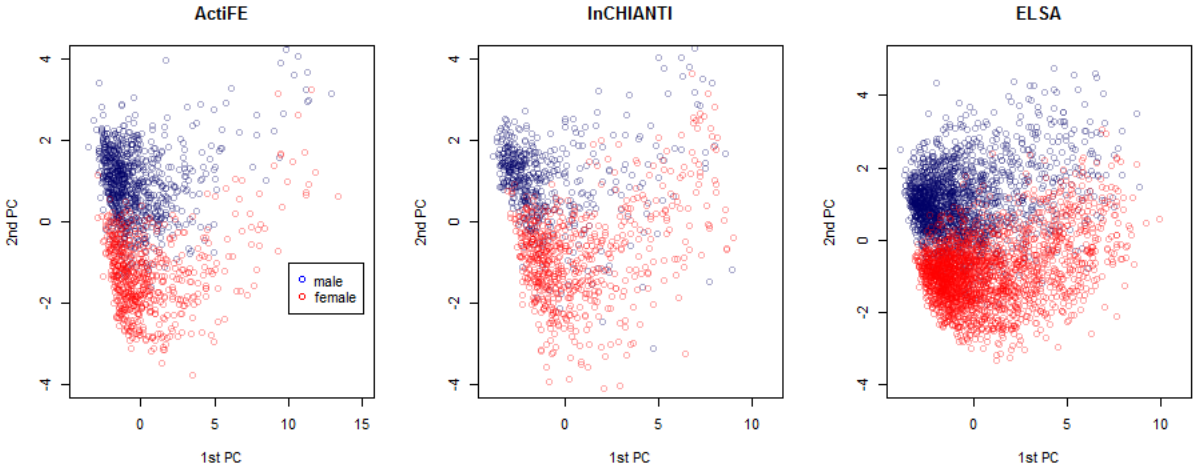

Figure 2. Scores on the first and second principal components of the observations from the three harmonized datasets. Men in blue, women in red.

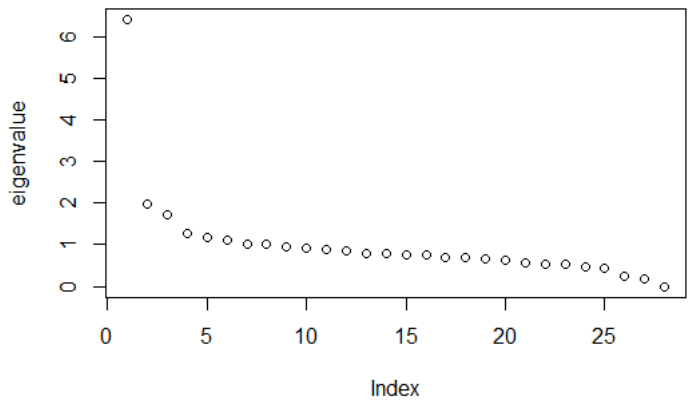

Figure 3. Eigenvalues of the correlation matrix of the three harmonized datasets stacked together

Table 4. Loadings of the first three PCs of the PCA performed on the three datasets stacked together.

	1st PC	2nd PC	3rd PC
Age	-0.221	0.028	0.029
Sex (women)	-0.111	0.177	-0.582
Gait speed	0.297	-0.158	-0.081
SPPB_BT score	0.291	-0.117	-0.122
SPPB_GST score	0.292	-0.103	-0.133
SPPB_CST score	0.279	-0.184	-0.104
SPPB score	0.352	-0.172	-0.145
History of falls (yes/no)	-0.119	-0.018	-0.154
History of falls (number)	-0.059	-0.045	-0.021
Grip strength	0.222	-0.211	0.418
Walking aid use	-0.256	0.021	0.166
Living alone	-0.093	0.192	-0.247
Urinary incontinence	-0.106	-0.250	-0.258
Diabetes	-0.055	-0.221	0.121
Parkinson disease	-0.043	-0.064	0.059
Arthritis or rheumatism	-0.112	-0.031	-0.264
Cognition impairment	-0.114	-0.117	0.069
History of stroke	-0.093	-0.043	0.128
Depression	-0.113	-0.143	-0.154
Pain	-0.122	-0.313	-0.154
Physical disability	-0.241	0.055	0.090
Instrumental disability	-0.272	-0.103	0.061
Fear of falling	-0.178	-0.200	-0.081
Dizziness	-0.162	-0.222	-0.119
Vision impairment	-0.077	-0.396	-0.053
Hearing impairment	-0.089	-0.099	0.170
Number of medication	-0.083	-0.478	-0.036
Physical activity limitations	-0.219	-0.147	0.111

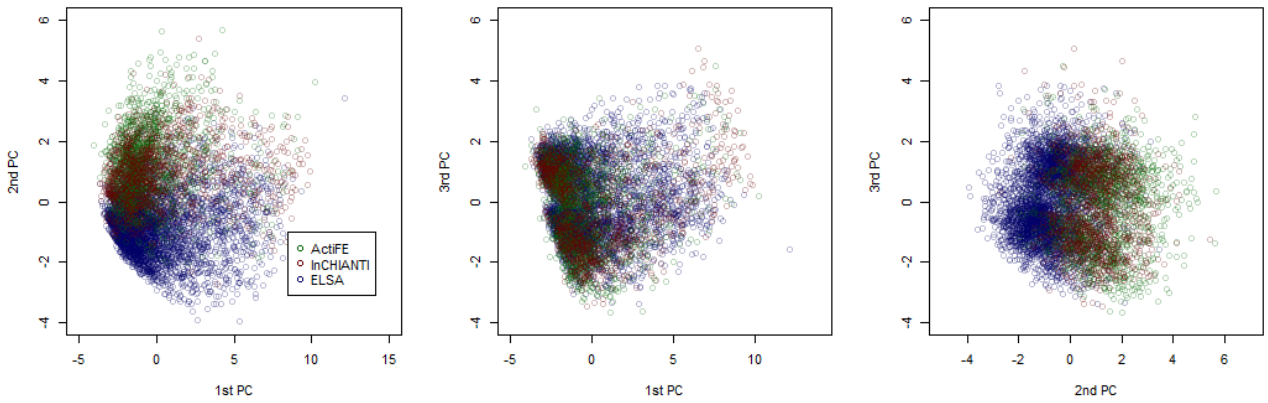

Figure 4. Scores on the first three principal components of the observations from the three harmonized datasets. PCA performed on the three datasets stacked together. Observations from ActiFE, InCHIANTI and ELSA are respectively in green, red and blue.

Associations of covariates with future falls

The association of the predictor target variables with the outcome target variables was quantified on the three populations with logistic and negative binomial regression models. In particular, logistic regressions were used for the dichotomous outcome variables “Prospective falls (yes/no)” and “Prospective multiple falls (yes/no)”, whereas negative binomial regressions were used for the count outcome variable “Prospective falls (number)”. Adjusted quantities refer to odds ratios or coefficients of negative binomial regressions where age and sex were included as covariates. Results are shown in Table 5, Table 6, and Table 7. Table 5 and Table 6 also show the unadjusted odds ratios reported in Deandrea et al. [1]. Adjusted odds ratios from that same study were not reported because other factors besides age and sex were used for adjusting. Weights were used on observations from the ELSA population in order to adjust for differential nonresponse rates to the nurse visit.

Inference on the single calculated quantities can be made using their specified 95 % confidence intervals. Statistical significance was more often reached in the ELSA population because of its large sample size. Comparisons between the calculated odds ratios and those reported in [1] is hindered by the fact that prediction intervals were not available from that meta-analysis [14].

Table 5. Odds ratios for any fall. 95 % confidence intervals are given in brackets.

	Deandrea	ActiFE		InCHIANTI		ELSA	
	Unadjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Age (years)	1.02 (1.01-1.03)	1.01 (0.99-1.02)		1.03 (1.01-1.06)		1.06 (1.05-1.07)	
Sex (women)	1.3 (1.18-1.42)	1.4 (1.12-1.75)		1.45 (1.05-1.99)		1.46 (1.26-1.69)	
Gait speed (m/s)		0.74 (0.49-1.11)	0.82 (0.52-1.3)	0.14 (0.07-0.27)	0.16 (0.08-0.35)	0.2 (0.14-0.27)	0.31 (0.22-0.44)
SPPB_BT score		0.89 (0.77-	0.91 (0.78-	0.74 (0.65-	0.78 (0.67-	0.7 (0.66-	0.77 (0.72-

		1.02)	1.05)	0.84)	0.9)	0.74)	0.82)
SPPB_GST score		0.96 (0.84-1.1)	0.99 (0.86-1.15)	0.69 (0.58-0.82)	0.74 (0.61-0.9)	0.65 (0.6-0.71)	0.72 (0.66-0.79)
SPPB_CST score		0.84 (0.76-0.93)	0.85 (0.76-0.94)	0.78 (0.69-0.88)	0.82 (0.71-0.93)	0.76 (0.72-0.8)	0.83 (0.79-0.88)
SPPB score		0.93 (0.89-0.98)	0.94 (0.89-0.99)	0.88 (0.84-0.93)	0.9 (0.84-0.95)	0.85 (0.82-0.87)	0.88 (0.85-0.91)
History of falls (yes/no)	2.77 (2.37-3.25)	1.57 (1.25-1.98)	1.54 (1.22-1.94)	1.94 (1.37-2.75)	1.75 (1.23-2.51)	3.96 (3.39-4.62)	3.6 (3.08-4.22)
History of falls (number)		1 (0.99-1.01)	1 (0.99-1.02)	1.48 (1.26-1.75)	1.42 (1.21-1.68)	1.48 (1.4-1.58)	1.43 (1.34-1.52)
Grip strength		0.98 (0.97-0.99)	0.99 (0.97-1)	0.98 (0.97-1)	0.99 (0.97-1.01)	0.96 (0.95-0.97)	0.97 (0.96-0.98)
Walking aid use	2.18 (1.79-2.65)	1.84 (0.7-4.8)	1.73 (0.65-4.57)	1.84 (1.09-3.09)	1.45 (0.83-2.56)	3 (2.36-3.83)	2.19 (1.7-2.83)
Living alone	1.33 (1.21-1.45)	1.06 (0.81-1.37)	0.95 (0.72-1.25)	1.36 (0.93-1.99)	1.17 (0.79-1.74)	1.6 (1.38-1.86)	1.24 (1.05-1.46)
Urinary incontinence	1.4 (1.26-1.57)	1.59 (1.26-2)	1.48 (1.16-1.88)	1.3 (0.94-1.78)	1.08 (0.77-1.52)	1.72 (1.43-2.07)	1.54 (1.28-1.87)
Diabetes	1.19 (1.08-1.31)	0.94 (0.67-1.31)	0.97 (0.69-1.35)	1.24 (0.8-1.91)	1.26 (0.81-1.95)	1.26 (1-1.58)	1.29 (1.02-1.64)
Parkinson disease	2.71 (1.08-6.84)	1.4 (0.63-3.15)	1.43 (0.63-3.22)	1.06 (0.29-3.97)	1.03 (0.27-3.86)	4.08 (1.68-9.92)	4.15 (1.69-10.19)
Arthritis or rheumatism	1.47 (1.28-1.70)	1.34 (1.07-1.68)	1.26 (1.01-1.59)	1.68 (1.22-2.32)	1.55 (1.11-2.16)	1.71 (1.48-1.99)	1.56 (1.34-1.82)
Cognition impairment	1.36 (1.12-1.65)	2.1 (0.52-8.44)	1.91 (0.47-7.74)	1.47 (0.92-2.35)	1.12 (0.68-1.85)	2.22 (0.92-5.39)	1.6 (0.64-4.03)
History of stroke	1.61 (1.31-1.98)	0.94 (0.57-1.56)	0.92 (0.55-1.54)	1.04 (0.5-2.15)	1.02 (0.49-2.12)	1.97 (1.5-2.59)	1.77 (1.34-2.34)
Depression	1.63 (1.36-1.94)	1.15 (0.8-1.66)	1.09 (0.75-1.59)	2.2 (1.51-3.19)	1.92 (1.31-2.83)	1.63 (1.29-2.05)	1.5 (1.18-1.9)
Pain	1.39 (1.19-1.62)	1.17 (0.93-1.48)	1.13 (0.9-1.43)	1.52 (0.91-2.54)	1.34 (0.79-2.26)	1.66 (1.43-1.92)	1.61 (1.38-1.87)
Physical disability	1.56 (1.22-1.99)	1.55 (0.85-2.84)	1.43 (0.77-2.64)	2.21 (1.21-4.02)	1.85 (0.99-3.44)	2.85 (2.38-3.4)	2.42 (2.02-2.91)
Instrumental disability	1.46 (1.20-1.77)	1.25 (0.9-1.76)	1.14 (0.79-1.63)	2.25 (1.6-3.18)	1.95 (1.3-2.92)	2.58 (2.12-3.15)	2 (1.63-2.46)
Fear of falling	1.55 (1.14-2.09)	1.42 (0.99-2.02)	1.34 (0.93-1.94)	1.86 (1.36-2.54)	1.59 (1.13-2.23)	3.46 (2.64-4.52)	2.71 (2.05-3.58)
Dizziness	1.80 (1.39-2.33)	1.11 (0.89-1.4)	1.05 (0.83-1.33)	0.98 (0.69-1.37)	0.94 (0.66-1.32)	2.62 (2.21-3.11)	2.23 (1.87-2.66)
Vision impairment	1.35 (1.18-1.54)	0.95 (0.7-1.3)	0.97 (0.7-1.33)	1.53 (1.09-2.15)	1.32 (0.91-1.92)	1.72 (1.46-2.02)	1.43 (1.21-1.69)
Hearing impairment	1.21 (1.05–1.39)	1.24 (0.96-1.6)	1.27 (0.97-1.67)	1.13 (0.79-1.61)	1.07 (0.73-1.55)	1.4 (1.2-1.65)	1.31 (1.11-1.55)
Number of medication	1.06 (1.04–1.08)	1.03 (0.99-1.07)	1.02 (0.98-1.07)	1.12 (1.04-1.2)	1.09 (1.01-1.17)	1.15 (1.07-1.24)	1.14 (1.06-1.23)
Physical activity limitations	1.20 (1.04–1.38)	1.07 (0.75-1.52)	1.01 (0.7-1.47)	2.45 (1.71-3.5)	2.1 (1.43-3.1)	2.26 (1.75-2.91)	1.71 (1.31-2.23)
Gait problems	2.06 (1.82–2.33)	1.02 (0.77-1.36)	0.94 (0.69-1.27)	2.73 (1.88-3.97)	2.43 (1.59-3.72)	2.09 (1.78-2.46)	1.67 (1.41-1.99)

Table 6. Odds ratios for multiple falls. 95 % confidence intervals are given in brackets.

	Deandrea	ActiFE		InCHIANTI		ELSA	
	Unadjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Age (years)	1.02 (1.01-1.03)	1.04 (1.02-1.07)		1.05 (1.02-1.08)		1.06 (1.04-1.07)	
Sex (women)	1.34 (1.12–1.60)	1.16 (0.84-1.61)		1.8 (1.11-2.92)		1.36 (1.12-1.65)	
Gait speed (m/s)		0.43 (0.23-0.79)	0.64 (0.32-1.26)	0.12 (0.05-0.29)	0.17 (0.06-0.52)	0.13 (0.09-0.2)	0.19 (0.12-0.3)
SPPB_BT score		0.76 (0.64-0.9)	0.82 (0.68-0.98)	0.72 (0.61-0.86)	0.78 (0.63-0.95)	0.7 (0.65-0.74)	0.75 (0.7-0.81)
SPPB_GST score		0.8 (0.68-0.94)	0.87 (0.73-1.04)	0.66 (0.53-0.83)	0.74 (0.57-0.95)	0.64 (0.58-0.71)	0.7 (0.63-0.78)
SPPB_CST score		0.76 (0.67-0.87)	0.81 (0.71-0.93)	0.73 (0.62-0.86)	0.79 (0.65-0.96)	0.69 (0.65-0.74)	0.75 (0.69-0.8)
SPPB score		0.88 (0.83-0.93)	0.9 (0.85-0.97)	0.87 (0.81-0.93)	0.9 (0.83-0.98)	0.82 (0.79-0.85)	0.84 (0.81-0.88)
History of falls (yes/no)	3.46 (2.85–4.22)	2.28 (1.64-3.17)	2.16 (1.55-3.01)	3.1 (1.94-4.96)	2.71 (1.68-4.37)	5.57 (4.53-6.84)	5.11 (4.14-6.31)
History of falls (number)		1.01 (0.99-1.02)	1.01 (0.99-1.02)	1.71 (1.41-2.06)	1.62 (1.35-1.96)	1.47 (1.38-1.56)	1.43 (1.35-1.52)
Grip strength		0.98 (0.97-1)	0.99 (0.96-1.01)	0.96 (0.94-0.99)	0.98 (0.95-1.02)	0.96 (0.95-0.97)	0.96 (0.94-0.97)
Walking aid use	3.09 (2.10–4.53)	4.14 (1.51-11.35)	3.28 (1.17-9.13)	1.86 (0.91-3.81)	1.31 (0.59-2.88)	3.58 (2.76-4.65)	2.68 (2.03-3.53)
Living alone	1.25 (1.10–1.43)	1.39 (0.96-1.99)	1.21 (0.83-1.78)	1.5 (0.88-2.57)	1.22 (0.7-2.11)	1.38 (1.14-1.68)	1.06 (0.86-1.32)
Urinary incontinence	1.67 (1.45–1.92)	1.99 (1.43-2.76)	1.94 (1.37-2.73)	1.55 (0.98-2.45)	1.2 (0.73-1.95)	2.08 (1.67-2.6)	1.9 (1.51-2.39)
Diabetes	1.28 (1.09–1.50)	1.63 (1.07-2.49)	1.61 (1.05-2.48)	1.11 (0.58-2.11)	1.13 (0.59-2.16)	1.3 (0.97-1.73)	1.33 (0.99-1.78)
Parkinson disease	2.84 (1.77–4.58)	3.67 (1.56-8.65)	3.24 (1.36-7.71)	0.97 (0.12-7.64)	0.94 (0.12-7.58)	4.48 (1.89-10.6)	4.41 (1.85-10.51)
Arthritis or rheumatism	1.57 (1.42–1.73)	1.36 (0.98-1.89)	1.33 (0.95-1.86)	1.63 (1.02-2.6)	1.43 (0.89-2.31)	1.7 (1.4-2.06)	1.56 (1.28-1.9)
Cognition impairment	1.56 (1.26–1.94)	7.62 (1.89-30.79)	5.93 (1.45-24.31)	2.01 (1.1-3.68)	1.43 (0.74-2.76)	1.98 (0.7-5.58)	1.36 (0.47-3.97)
History of stroke	1.79 (1.51–2.13)	1.19 (0.6-2.37)	1.07 (0.53-2.14)	1.07 (0.37-3.08)	1.05 (0.36-3.06)	1.96 (1.42-2.7)	1.72 (1.24-2.38)
Depression	1.86 (1.45–	1.58 (0.97-	1.4	2.08 (1.24-	1.66	2.24 (1.72-	2.13

	2.38)	2.56)	(0.86-2.29)	3.46)	(0.98-2.82)	2.93)	(1.62-2.8)
Pain	1.60 (1.44–1.78)	1.27 (0.9-1.78)	1.22 (0.87-1.72)	1.64 (0.73-3.65)	1.32 (0.58-3)	1.93 (1.59-2.34)	1.89 (1.55-2.3)
Physical disability	2.42 (1.80–3.26)	2.56 (1.27-5.17)	2.05 (1-4.2)	2.4 (1.12-5.15)	1.87 (0.84-4.16)	3.3 (2.68-4.06)	2.84 (2.29-3.52)
Instrumental disability	2.04 (1.41–2.95)	1.79 (1.16-2.77)	1.4 (0.87-2.24)	2.79 (1.74-4.47)	2.23 (1.27-3.93)	3.26 (2.6-4.08)	2.6 (2.05-3.29)
Fear of falling	2.51 (1.78–3.54)	1.81 (1.14-2.87)	1.52 (0.94-2.45)	2.29 (1.45-3.62)	1.81 (1.11-2.96)	3.64 (2.74-4.83)	2.9 (2.16-3.89)
Dizziness	2.28 (1.90–2.75)	1.8 (1.3-2.49)	1.66 (1.19-2.31)	0.98 (0.59-1.62)	0.93 (0.56-1.55)	3.2 (2.61-3.93)	2.76 (2.24-3.42)
Vision impairment	1.60 (1.28–2.00)	0.52 (0.35-0.77)	0.58 (0.39-0.88)	2.17 (1.28-3.67)	1.94 (1.09-3.44)	1.9 (1.55-2.32)	1.59 (1.29-1.96)
Hearing impairment	1.53 (1.33–1.76)	1.98 (1.4-2.79)	1.76 (1.22-2.55)	1.26 (0.76-2.09)	1.17 (0.68-2.01)	1.55 (1.27-1.9)	1.43 (1.16-1.77)
Number of medication	1.06 (1.04–1.08)	1.08 (1.02-1.13)	1.05 (1-1.11)	1.1 (0.99-1.22)	1.06 (0.95-1.18)	1.21 (1.1-1.32)	1.21 (1.1-1.33)
Physical activity limitations		1.37 (0.85-2.2)	1.11 (0.68-1.84)	3.08 (1.9-4.97)	2.46 (1.44-4.17)	2.8 (2.12-3.71)	2.17 (1.62-2.91)
Gait problems		1.48 (1.01-2.16)	1.16 (0.77-1.75)	3.28 (1.98-5.45)	2.73 (1.51-4.94)	2.23 (1.81-2.76)	1.82 (1.45-2.28)

Table 7. Coefficients from negative binomial regression models for number of falls. 95 % confidence intervals are given in brackets.

	ActiFE		InCHIANTI		ELSA	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Age (years)	0.02 (0.01-0.04)		0.03 (0.01-0.05)		0.02 (0-0.03)	
Sex (women)	0.09 (-0.12-0.31)		0.28 (-0.04-0.59)		0.16 (0-0.32)	
Gait speed (m/s)	-0.55 (-0.93--0.17)	-0.44 (-0.87--0.01)	-1.72 (-2.33--1.12)	-1.62 (-2.35--0.9)	-1.9 (-2.22--1.58)	-1.71 (-2.06--1.36)
SPPB_BT score	-0.26 (-0.39--0.13)	-0.23 (-0.37--0.1)	-0.29 (-0.41--0.16)	-0.26 (-0.4--0.12)	-0.35 (-0.41--0.29)	-0.35 (-0.41--0.28)
SPPB_GST score	-0.23 (-0.35--0.11)	-0.2 (-0.33--0.08)	-0.4 (-0.57--0.23)	-0.35 (-0.55--0.16)	-0.5 (-0.58--0.41)	-0.45 (-0.54--0.36)
SPPB_CST score	-0.2 (-0.29--0.11)	-0.19 (-0.28--0.1)	-0.24 (-0.36--0.12)	-0.2 (-0.34--0.07)	-0.42 (-0.47--0.36)	-0.41 (-0.47--0.36)
SPPB score	-0.11 (-0.15--0.06)	-0.1 (-0.15--0.05)	-0.12 (-0.17--0.07)	-0.11 (-0.17--0.05)	-0.21 (-0.24--0.18)	-0.2 (-0.23--0.16)
History of	0.71 (0.5-0.92)	0.69 (0.47-	0.88 (0.54-	0.82 (0.48-	1.64 (1.49-1.79)	1.63 (1.48-

falls (yes/no)		0.9)	1.21)	1.16)		1.78)
History of falls (number)	0.1 (0.09-0.11)	0.1 (0.09-0.11)	0.42 (0.28-0.55)	0.39 (0.26-0.53)	0.29 (0.28-0.31)	0.29 (0.27-0.3)
Grip strength	-0.02 (-0.03--0.01)	-0.02 (-0.04--0.01)	-0.02 (-0.04--0.01)	-0.02 (-0.04-0)	-0.02 (-0.03--0.02)	-0.04 (-0.06--0.03)
Walking aid use	0.84 (-0.02-1.7)	0.76 (-0.1-1.62)	0.51 (-0.02-1.05)	0.38 (-0.19-0.95)	1.56 (1.32-1.79)	1.6 (1.36-1.85)
Living alone	-0.01 (-0.26-0.24)	-0.12 (-0.39-0.14)	0.27 (-0.11-0.66)	0.12 (-0.28-0.51)	0.02 (-0.15-0.19)	-0.08 (-0.26-0.09)
Urinary incontinence	0.71 (0.5-0.92)	0.71 (0.48-0.93)	0.35 (0.03-0.66)	0.22 (-0.11-0.56)	0.66 (0.47-0.86)	0.65 (0.45-0.85)
Diabetes	0.43 (0.13-0.72)	0.43 (0.13-0.73)	0.09 (-0.35-0.54)	0.15 (-0.29-0.59)	0.45 (0.2-0.69)	0.43 (0.19-0.68)
Parkinson disease	0.85 (0.14-1.56)	0.82 (0.12-1.53)	-0.44 (-1.99-1.11)	-0.61 (-2.2-0.99)	1.78 (0.9-2.67)	1.76 (0.88-2.64)
Arthritis or rheumatism	0.22 (0.01-0.44)	0.2 (-0.01-0.42)	0.43 (0.11-0.76)	0.41 (0.09-0.74)	0.6 (0.44-0.75)	0.63 (0.47-0.79)
Cognition impairment	0.77 (-0.5-2.05)	0.55 (-0.73-1.83)	0.35 (-0.12-0.82)	0.05 (-0.45-0.55)	-0.02 (-1.07-1.03)	-0.09 (-1.13-0.96)
History of stroke	0.68 (0.25-1.11)	0.67 (0.24-1.1)	0.07 (-0.65-0.79)	0.1 (-0.6-0.81)	0.16 (-0.15-0.47)	0.13 (-0.17-0.44)
Depression	0.39 (0.05-0.72)	0.33 (-0.01-0.66)	0.61 (0.24-0.99)	0.52 (0.14-0.91)	0.61 (0.36-0.86)	0.6 (0.35-0.85)
Pain	0.31 (0.09-0.53)	0.31 (0.08-0.53)	0.33 (-0.16-0.82)	0.22 (-0.28-0.72)	0.77 (0.62-0.93)	0.83 (0.67-0.98)
Physical disability	1.22 (0.71-1.73)	1.19 (0.68-1.7)	0.83 (0.23-1.42)	0.75 (0.15-1.35)	1.33 (1.16-1.51)	1.36 (1.18-1.54)
Instrumental disability	0.57 (0.27-0.88)	0.49 (0.17-0.82)	0.75 (0.41-1.09)	0.64 (0.24-1.04)	1.49 (1.29-1.68)	1.51 (1.31-1.72)
Fear of falling	0.75 (0.44-1.07)	0.7 (0.37-1.02)	0.69 (0.39-1)	0.58 (0.25-0.92)	1.58 (1.31-1.84)	1.56 (1.29-1.82)
Dizziness	0.5 (0.29-0.71)	0.46 (0.24-0.68)	0 (-0.34-0.34)	-0.06 (-0.4-0.28)	1.38 (1.21-1.55)	1.37 (1.2-1.55)
Vision impairment	-0.31 (-0.59--0.02)	-0.23 (-0.53-0.06)	0.47 (0.13-0.81)	0.36 (-0.01-0.73)	0.39 (0.21-0.56)	0.33 (0.15-0.51)
Hearing impairment	0.39 (0.15-0.63)	0.34 (0.08-0.59)	0.24 (-0.11-0.59)	0.2 (-0.16-0.57)	0.27 (0.09-0.44)	0.24 (0.07-0.42)
Number of medication	0.09 (0.06-0.13)	0.08 (0.05-0.12)	0.1 (0.02-0.17)	0.08 (0-0.15)	0.25 (0.17-0.33)	0.26 (0.18-0.34)
Physical activity limitations	0.53 (0.22-0.85)	0.47 (0.14-0.8)	0.8 (0.45-1.15)	0.68 (0.3-1.06)	1.21 (0.95-1.47)	1.18 (0.91-1.45)
Gait problems	0.44 (0.19-0.7)	0.37 (0.09-0.64)	0.97 (0.61-1.33)	0.9 (0.49-1.32)	0.97 (0.8-1.13)	0.87 (0.7-1.05)

Optimum reference

The performance of a predictive model of interest is often compared to a reference model that we intend to outperform (e.g. a dummy or state-of-the-art model) or an 'optimal' model that we intend to approach. For example, we recall the definition of the skill score (SS) of a forecast (see e.g. [15]):

$$SS = \frac{score\ of\ the\ forecast - score\ of\ the\ reference\ forecast}{score\ of\ the\ optimal\ forecast - score\ of\ the\ reference\ forecast}$$

For the AUC, natural lower and upper reference values may be 0.5 and 1. For the BS, they are UNC and 0. However, the BS can reach zero only if the model is deterministic. Similarly, the AUC of a probabilistic model is expected to be lower than 1.

In this section, we have taken as optimal reference score the score that our model of interest (FRAT-up) would attain if evaluated on the same subjects and under the hypothesis of perfect calibration of the model.

We call $\bar{x}$ the observed n-by-p matrix of covariates, with n number of subjects of the sample and p number of covariates measured on each subject. We call $\bar{y}$ the n-by-1 vector of 1's and 0's that is the target of our prediction. We call f our model of interest and $\hat{y} = f(\bar{x})$ our probabilistic predictions. In particular, the i-th component of $\hat{y}$ is the probability, assigned by the model f, for the i-th component of $\bar{y}$ to be 1. We call $S(\bar{y}, f(\bar{x}))$ the score assigned to our model f. For example, S may be the AUC or the BS.

We propose $S(Y, f(\bar{x}))$, with $Y|X = \bar{x} \sim f(\bar{x})$ as the optimum reference score for a fixed dataset of covariates $\bar{x}$. As Y is a random quantity, also $S(Y, f(\bar{x}))$ is.

The notation $Y|X = \bar{x} \sim f(\bar{x})$ means that the i-th component of Y follows a Bernoulli distribution with parameter equal to the i-th component of $f(\bar{x})$, for every $i = 1,2, \dots, n$.

In order to estimate the distribution of $S(Y, f(\bar{x}))$, we draw a vector y_k from a distribution described by $f(\bar{x})$. We then calculate $s_k = S(y_k, f(\bar{x}))$, and repeat this procedure for $k = 1,2, \dots, K$, with $K = 2000$. The s_k's are then distributed as $S(Y, f(\bar{x}))$, for $Y|X = \bar{x} \sim f(\bar{x})$.

In Figure 5 the actual values $S(\bar{y}, f(\bar{x}))$ are compared to the distributions of the optimum quantities $S(Y, f(\bar{x}))$, $Y|X = \bar{x} \sim f(\bar{x})$, for the three datasets and for S being AUC and GRES. The optimal median AUCs are ordered as the actual values, with the optimal median AUC for ELSA being the highest and the optimal median AUC for ActiFE being the lowest. However, they take less extreme values and the distributions are largely overlapped. The same is valid for GRES, though the ranking between the median GRES on InCHIANTI and ELSA is inverted.

These results suggest that the differences in the distributions of the risk factors among the three populations only partially explain the differences in the performance that FRAT-up attains in the three dataset.

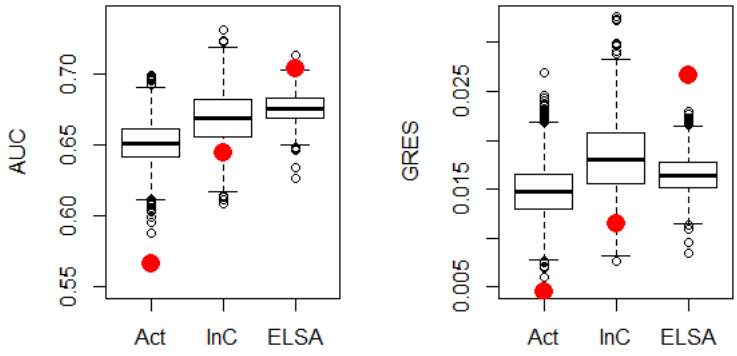

Figure 5. Comparison between optimal and actual values of AUC and GRES attained by FRAT-up on the three datasets. The actual values $S(\bar{y}, f(\bar{x}))$ are marked in red. S is the AUC (left panel) and the GRES (right panel). The distributions of the optimal values of AUC and GRES are shown as boxplots computed from the s_k's, derived as explained in the text.

Sensitivity analysis

The scope of the sensitivity analysis here presented is to identify FRAT-up most influential variables.

The strategy is to compare the performance of FRAT-up with the performance of some of its alterations on the three datasets. The scores that are used to quantify the performance are the AUC, and the Brier skill score (BSS). The BSS is defined as $BSS = 1 - BS/BS_{ref}$, where BS is the Brier score and BS_{ref} is the Brier score of a classifier that assigns each subject a risk equal to the prevalence in the population.

The alterations are produced via variable removal and variable permutation.

- Sensitivity analysis through variable removal. Values of one risk factor are set to not available. A score of the performance of the model is calculated. This procedure is repeated for each risk factor.
- Sensitivity analysis through variable permutation. Values of one risk factor are permuted 50 times across the samples in the dataset. For each permutation, a score of the performance of the model is calculated. A mean score is obtained averaging across permutations. The procedure is repeated for each risk factor.

Results are shown in Table 8, Table 9 and Figure 6.

Table 8. Sensitivity analysis through variable removal. Δ indicates the increase in the performance score. Negative values indicate performance decrease in the altered models.

	Δ AUC			Δ BSS		
	ActiFE	InCHIANTI	ELSA	ActiFE	InCHIANTI	ELSA
Age	0.0009	-0.0013	-0.0028	0.0026	-0.0011	-0.0021
Female sex	-0.0034	0.0004	-0.0006	-0.0013	0.0025	-0.0011
Living alone	0.0002	-0.0008	-0.0019	0.0014	-0.0025	-0.0018
History of falls	-0.0124	-0.0059	-0.0404	-0.0019	-0.0109	-0.0365
Physical activity limitation	0.0005	-0.0009	0.0003	0.0012	-0.0041	0.0018
Physical disability	0.0004	0.0002	-0.0006	0.0010	-0.0012	-0.0028
Instrumental disability	0.0007	-0.0012	0.0009	0.0021	0.0020	-0.0004
Walking aid use	0.0006	-0.0006	0.0006	0.0016	-0.0029	0.0003
Cognition impairment	-0.0001	-0.0005	-0.0001	0.0003	-0.0016	0.0020
Depression	0.0009	-0.0067	0.0020	0.0030	-0.0017	0.0011
History of stroke	0.0012	0.0009	-0.0003	0.0018	-0.0011	0.0001
Urinary incontinence	-0.0066	0.0035	-0.0012	-0.0019	0.0092	-0.0011
Rheumatic disease	-0.0053	-0.0055	0.0005	-0.0010	-0.0075	-0.0009
Dizziness	0.0041	0.0063	-0.0010	0.0076	0.0149	-0.0045
Diabetes	0.0007	-0.0001	0.0006	0.0010	0.0004	0.0002
Poor self-perceived health status	0.0004	-0.0009	NA	0.0038	-0.0045	NA
Pain	0.0006	0.0006	0.0021	0.0027	0.0166	-0.0017
Fear of falling	-0.0002	-0.0041	-0.0009	0.0029	0.0001	0.0017
Parkinson disease	-0.0012	0.0020	-0.0015	0.0001	0.0022	-0.0014
Sedatives	0.0004	-0.0001	NA	0.0006	-0.0015	NA
Antihypertensives	0.0004	-0.0014	NA	0.0010	0.0005	NA
Antiepileptics	-0.0022	-0.0003	NA	-0.0012	0.0001	NA
Gait problems	0.0080	-0.0080	0.0031	0.0093	-0.0149	0.0003
Vision impairment	-0.0008	-0.0012	0.0006	-0.0033	0.0043	-0.0006
Hearing impairment	-0.0002	0.0001	-0.0001	0.0006	-0.0006	0.0002
Number of medications	0.0007	-0.0008	0.0008	0.0022	0.0007	0.0020

Table 9. Sensitivity analysis through variable permutation. Δ indicates the increase in the performance score. Negative values indicate performance decrease in the altered models.

	Δ AUC			Δ BSS		
	ActiFE	InCHIANTI	ELSA	ActiFE	InCHIANTI	ELSA
Age	0.0009	-0.0021	-0.0037	0.0023	-0.0015	-0.0032
Female sex	-0.0032	0.0002	-0.0013	-0.0015	0.0009	-0.0010
Living alone	0.0002	-0.0011	-0.0027	0.0011	-0.0008	-0.0022
History of falls	-0.0161	-0.0106	-0.0547	-0.0099	-0.0105	-0.0453
Physical activity limitation	0.0005	-0.0014	0.0001	0.0009	-0.0014	0.0000
Physical disability	0.0004	-0.0003	-0.0019	0.0008	-0.0016	-0.0025
Instrumental disability	0.0004	-0.0017	0.0000	0.0017	-0.0021	-0.0005
Walking aid use	0.0006	-0.0016	-0.0021	0.0011	-0.0022	-0.0029
Cognition impairment	-0.0001	-0.0012	0.0000	-0.0001	-0.0022	0.0000
Depression	-0.0003	-0.0079	0.0007	0.0013	-0.0052	-0.0001
History of stroke	0.0009	0.0003	-0.0015	0.0010	0.0011	-0.0017
Urinary incontinence	-0.0070	0.0025	-0.0031	-0.0029	0.0020	-0.0026
Rheumatic disease	-0.0065	-0.0071	-0.0024	-0.0035	-0.0035	-0.0028
Dizziness	0.0026	0.0037	-0.0055	0.0050	0.0059	-0.0066
Diabetes	0.0004	-0.0006	0.0001	0.0006	-0.0006	-0.0002
Poor self-perceived health status	-0.0004	-0.0013	NA	0.0023	-0.0014	NA
Pain	-0.0009	0.0010	-0.0007	0.0019	0.0026	-0.0015
Fear of falling	-0.0005	-0.0053	-0.0016	0.0019	-0.0043	-0.0021
Parkinson disease	-0.0013	0.0013	-0.0022	-0.0004	0.0009	-0.0018
Sedatives	0.0004	-0.0002	NA	0.0004	-0.0006	NA
Antihypertensives	0.0004	-0.0015	NA	0.0013	-0.0008	NA
Antiepileptics	-0.0020	-0.0003	NA	-0.0011	-0.0001	NA
Gait problems	0.0070	-0.0105	-0.0037	0.0071	-0.0100	-0.0051
Vision impairment	-0.0008	-0.0017	-0.0004	-0.0014	-0.0007	-0.0008
Hearing impairment	-0.0003	0.0001	-0.0004	0.0003	0.0001	-0.0004
Number of medications	0.0005	-0.0005	0.0007	0.0023	0.0001	0.0004

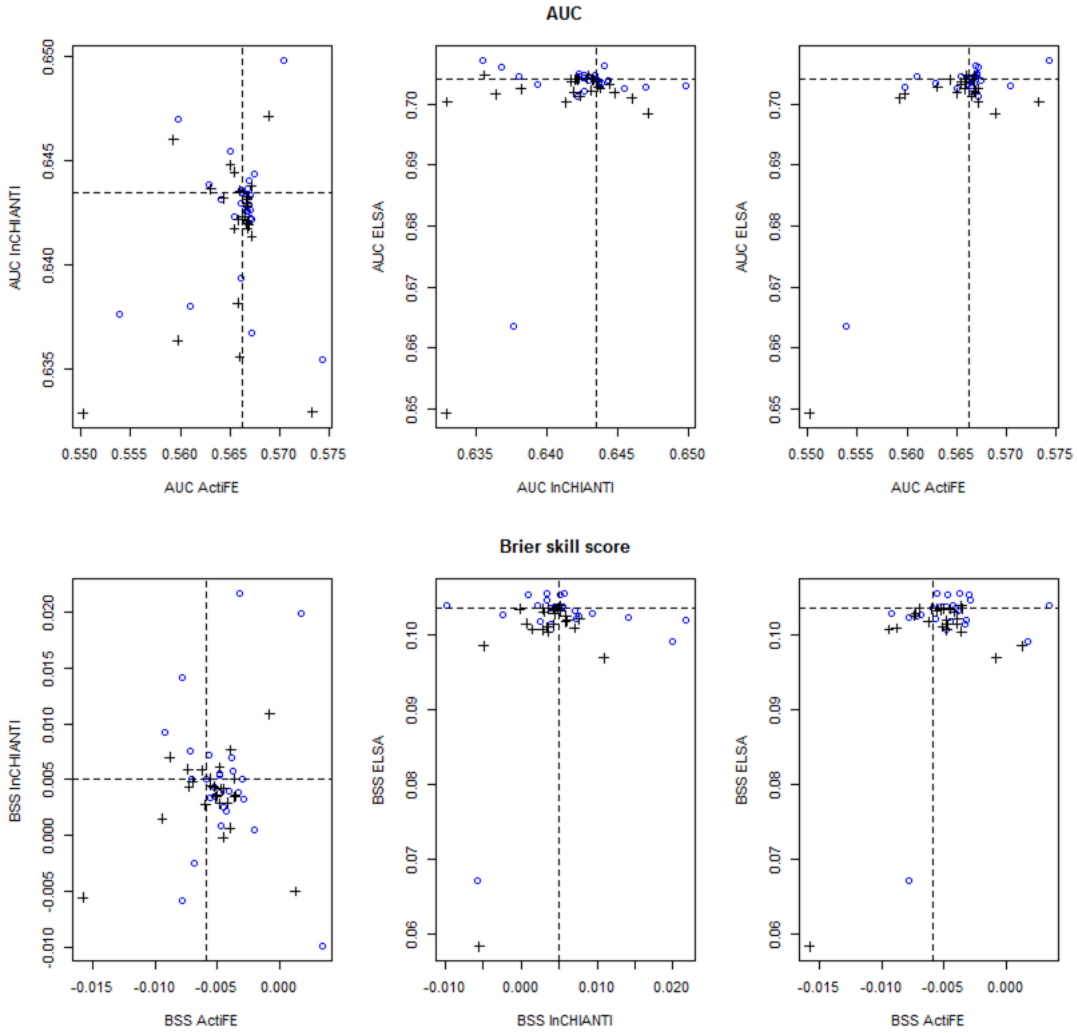

Figure 6. AUC and Brier skill score of the altered models on the three datasets. Each circle or cross represents a risk factor. Circles are relative to model alteration through variable removal. Crosses are relative to model alteration through variable permutation. Horizontal and vertical dotted lines indicate performance of the unaltered FRAT-up.

References

1. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk Factors for Falls in Community-dwelling Older People: A Systematic Review and Meta-analysis. Epidemiology 2010;21 (5):658–668.

2. Istituto nazionale di statistica. Statistiche ISTAT. [date unknown];.

3. Capon A, Lallo D Di, Mastromattei A, Pavoni N, Simeoni S. Incidence and risk factors for accidental falls among general practice elderly patients in Latina, Central Italy. 2007;31 (4):204–211.

4. Tromp a M, Pluijm SM, Smit JH, Deeg DJ, Bouter LM, Lips P. Fall-risk screening test: a prospective study on predictors for falls in community-dwelling elderly. J. Clin. Epidemiol. 2001 Aug;54 (8):837–44. http://www.ncbi.nlm.nih.gov/pubmed/11470394.
.

5. Stone KL, Ancoli-israel S, Blackwell T, Ensrud KE, Cauley JA, Redline S, Hillier TA, Schneider J, Claman D, Cummings SR. Actigraphy-Measured Sleep Characteristics and Risk of Falls in Older Women. 2008;168 (16):1768–1775.

6. Gaßmann KG, Rupprecht R, Freiberger E. Predictors for occasional and recurrent falls in community-dwelling older people. Z. Gerontol. Geriatr. 2009;42:3–10.

7. Chu L, Chi I, Chiu A. Incidence and predictors of falls in the Chinese elderly. Ann Acad Med Singapore 2005;60–72.

8. Covinsky KE, Kahana E, Kahana B, Kercher K, Schumacher JG, Justice AC. History and mobility exam index to identify community-dwelling elderly persons at risk of falling. J. Gerontol. A. Biol. Sci. Med. Sci. 2001 Apr;56 (4):M253–9. http://www.ncbi.nlm.nih.gov/pubmed/11283200. Accessed 2014-4-1.

9. World Health Organization Department of Ageing and Life Course. WHO global report on falls prevention in older age. Geneva, Switzerland: World Health Organization; 2008.

10. National Health Information Council. National Population Health Survey: Health Institutions Component, Longitudinal (NPHS). 2007;.

11. Pluijm SMF, Smit JH, Tromp E a M, Stel VS, Deeg DJH, Bouter LM, Lips P. A risk profile for identifying community-dwelling elderly with a high risk of recurrent falling: results of a 3-year prospective study. Osteoporos. Int. 2006 Jan;17 (3):417–25. http://www.ncbi.nlm.nih.gov/pubmed/16416256. Accessed 2013-6-7.

12. Townsend N, Bhatnagar P, Wickramasinghe K, Scarborough P, Foster C, Rayner M. Physical activity statistics 2012. 2012.

13. Clough-Gorr KM, Erpen T, Gillmann G, von Renteln-Kruse W, Iliffe S, Beck JC, Stuck AE. Multidimensional Geriatric Assessment: Back to the Future Preclinical Disability as a Risk Factor for Falls in Community-Dwelling Older Adults. Journals Gerontol. Ser. A Biol. Sci. Med. Sci. 2008 Mar;63 (3):314–320.

14. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011 Jan;342 (feb10_2):d549. http://www.bmj.com/content/342/bmj.d549. Accessed 2014-8-31.

15. Gneiting T, Raftery AE. Strictly Proper Scoring Rules, Prediction, and Estimation. J. Am. Stat. Assoc. 2007 Mar;102 (477):359–378.

http://www.tandfonline.com/doi/abs/10.1198/016214506000001437. Accessed 2013-8-15.

Appendix 3. ActiFE

This appendix contains a description of how the variables in the ActiFE dataset have been processed to validate FRAT-up.

The variables of the ActiFE dataset are called 'source variables'. The variables that are needed in FRAT-up are called 'target variables'. Among the target variables we distinguish the predictor target variables and an outcome target variable. In Chapter 4, predictor target variables are also called 'risk factors' and the source variables used to derive these target variables are called 'risk factor estimators'.

For the sake of clarity and brevity, the documentation does not take into account in details how missing values (not applicable, refusal, unknown…) are coded and handled. To have this kind of information, please refer to the original documentation of the database and contact us for the R scripts.

Outcome target variable

Prospective falls (yes/no)

Source variable: "sturz".

Description of "sturz". Whether the subject reports at least one fall in their fall calendar. Possible values: 0 (=no fall), 1 (=at least one fall). The fall calendar covers about 12 months after the baseline assessment, but the exact time coverage slightly changes among subjects.

Conversion. Prospective falls (yes/no) = sturz

Possible values for "Prospective falls (yes/no)": 0, 1.

Predictor target variables

Age

Source variable: "alter".

Description of "alter". Age at baseline in years. Range of values [65.3, 91.4].

Conversion. Age = alter.

Sex

Source variable: "sex".

Description of "sex". Gender. Possible values: 1 (=male), 2 (=female).

Conversion. Sex = sex-1.

Living situation

Source variable: "IV1N022".

Description of "IV1N022". Question "Are there any persons you know living in your house?". Possible values: 0 (=no), 1 (=yes).

Conversion. Living situation = 1 - IV1N022.

History of falls (yes/no)

Source variable: “sturz12”.

Description of “sturz12”. Label in the dataset: fall in the last 12 months (self-report). Possible values: 0 (=no), 1 (=yes).

Conversion. History of falls (yes/no) = sturz12.

History of falls (number)

Source variable: “sturz12_anz”.

Description of “sturz12_anz”. Label in the dataset: number of falls in the last 12 months (self-report).

Conversion. History of falls (number) = sturz12_anz.

Physical activity limitation

Source variable: “walktime_daily”.

Description of “walktime_daily”. Label in the dataset: mean time of walking per day (min) from ActivePal`s.

Conversion. If walktime_daily <60 { Physical activity limitation = 1} otherwise { Physical activity limitation = 0 }.

Physical disability

Source variable: “adl_aid”.

Description of “adl_aid”. Label in the dataset: number of activities which require aid (ADL). Possible values: 0,1,2,3,4,5. Subjects are asked on the following activities:

- Taking a shower or bath
- Getting dressed and undressed
- Sitting on and getting up from a chair
- Going up and down a staircase with 15 steps without making a break
- Making a 5-min walk without stopping

Conversion rule. If adl_aid>=1 {Physical disability = 1}, otherwise { Physical disability = 0}.

Instrumental disability

Source variable “iadl_aid”.

Description of “iadl_aid”. Label in the dataset: number of activities which require aid (IADL). Possible values: 0,1,2,3,4,5. Subjects are asked on the following activities:

- Using their own means of transport (e.g. bicycle, car) or public transport (e.g. bus, train)
- Doing independently light housework (making beds, washing dishes…)
- Managing independently the drugs to take
- Going shopping independently
- Cutting their own toenails

Conversion rule. If iadl_aid>=1 {Instrumental disability = 1}, otherwise { Instrumental disability = 0}.

Cognitive impairment

Source variable: “mmse”.

Description of “mmse”. Label in the dataset: ‘Mini-Mental-State-Examination (MMSE)’. Possible values: 0,1,2,…,30.

Conversion rule. If mmse<=20 {Cognitive impairment = 1}, otherwise { Cognitive impairment = 0}.

Depression

Source variable: “hads_d”.

Description of “hads_d”. “hads_d” is the depression subscore of the Hospital Anxiety and Depression Scale (HADS). Possible values: 0,1,2,…,21.

Conversion rule. If hads_d>=8 {Depression = 1}, otherwise {Depression = 0}.

History of stroke

Source variable: “stroke”

Description of “stroke”. Question: ‘Has a doctor ever told you that you have or had any of these diseases?...Stroke (no TIA)’. Possible values: 0 (=no), 1 (=yes and I am still impaired / yes but it does not affect me any more).

Conversion rule. History of stroke = stroke.

Urinary incontinence

Source variable: “IV3N778”.

Description of “IV3N778”. Question ‘When you lose involuntarily urine?’.

Conversion rule. If IV3N778=0 (=never - no urine leakage) {Urinary incontinence = 1}, otherwise { Urinary incontinence = 0}. See the R code for knowing more.

Rheumatic disease

Source variables: “arthritis”, “rheuma”.

Description of “arthritis”. Question ‘Has a doctor ever told you that you have or had any of these diseases?... Arthrosis / arthritis’. Possible values: 0 (=no), 1 (=yes and I am still impaired / yes but it does not affect me any more).

Description of “rheuma”. Question ‘Has a doctor ever told you that you have or had any of these diseases?... Rheumatic diseases? As classical rheumatoid or lupus (not arthrosis / arthritis)’. Possible values: 0 (=no), 1 (=yes and I am still impaired / yes but it does not affect me any more).

Conversion rule. If arthritis==1 or rheuma==1 { Rheumatic disease = 1}, otherwise { Rheumatic disease = 0}

Dizziness and vertigo

Source variable: “IV3N791”.

Description of “IV3N791”. Question ‘Do you suffer of vertigo’. Possible values: 0 (=never),1-4 (=rarely-constantly).

Conversion rule. If IV3N791==0 {Dizziness and vertigo = 0}, otherwise { Dizziness and vertigo = 1}.

Diabetes

Source variable: “diab”

Description of “diab”. Question: ‘Has a doctor ever told you that you have or had any of these diseases?...Diabetes/diabetes mellitus’. Possible values: 0 (=no), 1 (=yes and I am still impaired / yes but it does not affect me any more).

Conversion rule. Diabetes = diab.

Self-perceived health status

Source variable: “sah”.

Description of “sah”. Question ‘How would you describe your health in general?’. Possible values: 0 (=less than well, bad), 1(=excellent, very good, good).

Conversion rule. Self-perceived health status = sah.

Pain

Source variables: “pain”, “backpain”.

Description of “pain”. Question ‘Do you have chronic pain (not back pain, osteoarthritis, arthritis)?’. Possible values: 0 (=no), 1 (=yes and I am still impaired / yes but it does not affect me any more).

Description of “backpain”. Question ‘Do you have back pain?’. Possible values: 0 (=no), 1 (=yes and I am still impaired / yes but it does not affect me any more).

Conversion rule. If pain==1 or backpain==1 { Pain = 1}, otherwise { Pain = 0}.

Fear of falling

Source variable: “fesi”.

Description of “fesi”. Label in the dataset ‘Short FES-I’. Possible values: 7,8,9,…,28. Activities on which subjects are asked are:

- Dressing and undressing
- Taking a shower
- Getting up from and sitting down to a chair
- Going upstairs or downstairs
- Grasping something above the head or on the floor
- Moving up or down a slope
- Going to an event (e.g. a family reunion, an association meeting or a service)

Conversion rule. If fesi>=11 {Fear of falling = 1}, otherwise {Fear of falling = 0}.

Parkinson disease

Source variable: “ATC”.

Description of “ATC”. ATC codes of the medications used by the subject.

Conversion rule. If at least one of the ATC codes is in the list ('N04BA02', 'N04BA03', 'N04BA01', 'N04BC01', 'N04BC06', 'N04BC10', 'N04BC02', 'N04BC05', 'N04BC04', 'N04BC09', 'N04BD02', 'N04BD01', 'N04BX02', 'N04AA02', 'N04AA11', 'N04AA03', 'N04AA04', 'N04AA01', 'N04BB01', 'N04BX03', 'N04BX01', 'N04BC07') {Parkinson = 1}, otherwise {Parkinson = 0}.

Number of medications

Source variables: “ATC”.

Description of “ATC”. ATC codes of the medications used by the subject.

Conversion rule. Number of medications = number of unique ATC codes associated to a given subject.

Sedatives

Source variables: “ATC”.

Description of “ATC”. ATC codes of the medications used by the subject.

Conversion rule. If at least one of the ATC codes begins with ‘N05C’ {Sedatives = 1}, otherwise {Sedatives = 0}.

Antihypertensives

Source variables: “ATC”.

Description of “ATC”. ATC codes of the medications used by the subject.

Conversion rule. If at least one of the ATC codes is in the list (“C09CA[01-04,06-08]”, “C09DA[01-04,06,08”, “C03AA03”, “C03BA04”, “C03BA11”, “C03CA[01,03,04]”, “C03EA01”, “C03EB01”, “QC03CA0”, “C09AA”, “C09AA[01-13,15]”, “C09BA[01-09,12]”, “C08CA[01-09,11-13]”, “C08DA01”, “C08DB01”, “C07AA05”, “C07AA07”, “C07AA12”, “C07AB[02,03,07,08,12]”, “C07AG02”, “C07CA02”, “C07CB02”, “C07CB03”, “C02AC01”, “C02CA04”, “C02CA05”, “C02CA49”, “C02KA49”, “C02KD01”, “C02LA51”, “C03DA[01-03]” { Antihypertensives = 1}, otherwise { Antihypertensives = 0}.

Antiepileptics

Source variables: “ATC”.

Description of “ATC”. ATC codes of the medications used by the subject.

Conversion rule. If at least one of the ATC codes is in the list (“N03AA02”, “N03AA03”, “N03AB02”, “N03AB52”, “N03AF01”, “N03AF02”, “N03AG01”, “N03AX12”, “N03AX16”) { Antiepileptics = 1}, otherwise { Antiepileptics = 0}.

Grip strength

Source variable: “handkraft”.

Description of “handkraft”. Grip strength was measured two times for each hand, with arm at 90° on the table. “handkraft” is the maximum between the mean grip strength measured on the right hand, and the mean grip strength measured on the left hand.

Conversion rule. Grip strength = handcraft.

Gait speed

Source variable: “speed”.

Description of “speed”. Time for walking a giving distance (4 m on some subjects, 3 m on others) at self-select gait speed was measured twice. “speed” was computed as the distance (4 m or 3 m) divided by the mean of the two recorded times.

Conversion rule. Gait speed = speed.

SPPB balance

Source variable: “sppb_bt”.

Description of “sppb_bt”. Label in the dataset: ‘SPPB-Balance Test’. Possible values: 0,1,2,3,4. It was computed from:

- “iv1n349” (=balance test performed; 0=no, 1=yes),
- “iv1n350” (=Side by side (feet together) held more than 10 sec?; 1=yes, 0=no, 2=not carried out),
- “iv1n351” (=Semi-tandem stand held for more than 10 sec; 1=yes, 0=no, 2=not carried out),
- “iv1n352” (=Tandem stand held for more than 10; 1=yes, 0=no, 2=not carried out),
- “iv1n353” (=Tandem stand, held time in sec.)

according to the following algorithm:

if iv1n349=0 or iv1n350=0 or iv1n350=2 then sppb_bt = 0;

if iv1n350=1 and (iv1n351=0 or iv1n351=2) then sppb_bt = 1;

if iv1n350=1 and iv1n351=1 and (iv1n352=2 or iv1n352=0 or iv1n353 < 3) then sppb_bt = 2;

if iv1n350=1 and iv1n351=1 and iv1n353 >= 3 and iv1n353 < 10 then sppb_bt = 3;

if iv1n352 = 1 then sppb_bt = 4;

if iv1n352=.M and iv1n353=.M then sppb_bt = .;

if iv1n352=0 and iv1n353 in (.M,.A) then sppb_bt = .;

Conversion rule. SPPB balance = sppb_bt

SPPB gait

Source variable: “sppb_gst”.

Description of “sppb_gst”. Label in the dataset: ‘SPPB-Gait Speed Test’. Possible values: 0,1,2,3,4. Time for walking a giving distance (4 m on some subjects, 3 m on others) at self-select gait speed was measured twice. “gaitspeed” was computed as the maximum of the two velocities measured from the two trials. “sppb_gst” was computed from “gaitspeed” according to the following algorithm:

if the walking test is not performed then sppb_gst = 0;

if gaitspeed <= 0.43 and gaitspeed ne . then sppb_gst = 1;

if gaitspeed > 0.43 and gaitspeed <= 0.60 then sppb_gst = 2;

if gaitspeed > 0.60 and gaitspeed <= 0.77 then sppb_gst = 3;

if gaitspeed > 0.77 then sppb_gst = 4;

Conversion. SPPB gait = sppb_gst.

SPPB chair standing

Source variable: “sppb_cst”.

Description of “sppb_cst”. Label in the dataset: ‘SPPB-Chair Stand Test’. Possible values: 0,1,2,3,4. “iv1n367” is the time for standing five time from the chair. “sppb_cst” is calculated according to the algorithm:

if test not performed then sppb_cst = 0;

if iv1n367 >= 16.7 then sppb_cst = 1;

if iv1n367 >= 13.6 and iv1n367 < 16.7 then sppb_cst = 2;

if iv1n367 >= 11.2 and iv1n367 < 13.6 then sppb_cst = 3;

if iv1n367 < 11.2 and iv1n367 not in (.,.A,.M) then sppb_cst = 4;

Conversion. SPPB chair standing = sppb_cst.

SPPB

Source variable: “sppb”.

Description of “sppb”. Label in dataset: ‘SPPB’. Possible values: 0,1,2,…,12. “sppb” is the sum of “sppb_bt”, “sppb_gst” and “sppb_cst”, whose description is given above.

Conversion. SPPB = sppb.

Gait problems

Source variable: “sppb_gst”.

Description of “sppb_gst”. It is given above at the paragraph “SPPB gait”.

Conversion. If sppb_gst<4 {Gait problems = 1}, otherwise { Gait problems = 0}.

Walking aid use

Source variable: “IV1N359”.

Description of “IV1N359”. Walking aid used during the walking test. Possible values: 1 (= no walking aid), 2,3,4 (= different walking aids).

Conversion. If IV1N359==1 {Walking aid use = 0}, otherwise { Walking aid use = 1}.

Vision impairment

Source variable: “IV3N793”.

Description of “IV3N793”. Scores according to Jaeger, 35 cm distance.

Conversion. If IV3N793 < 5 {Vision impairment = 1}, otherwise { Vision impairment = 0}.

Hearing impairment

Source variable: “IV3N792”.

Description of “IV3N792”. How well the subject hears, rated by the interviewer. Possible values: 4 (=very good), 3 (=good), 2 (=fair), 1 (=poor).

Conversion. If IV3N792== 1 or 2 { Hearing impairment = 1}, otherwise { Hearing impairment = 0}.

Appendix 3. ELSA

This appendix contains a description of how the variables in the ELSA dataset have been processed to validate FRAT-up.

The variables of the ELSA dataset are called 'source variables'. The variables that are needed in FRAT-up are called 'target variables'. Among the target variables we distinguish the predictor target variables and an outcome target variable. In Chapter 4, predictor target variables are also called 'risk factors' and the source variables used to derive these target variables are called 'risk factor estimators'.

For the sake of clarity and brevity, the documentation does not take into account in details how missing values (not applicable, refusal, unknown…) are coded and handled. To have this kind of information, please refer to the original documentation of the database and contact us for the R scripts.

The outcome variable is taken from ELSA wave 3. Predictor variables are taken from ELSA wave 2. Wave 2 is chosen as baseline assessment because it is the first wave during which a nurse visit was carried out. Wave 3 was carried out about 2 years after wave 2.

Outcome target variable

Prospective falls (yes/no)

Source variable: "hefla"

Description of "hefla". Whether the subject has ever fallen down since the last interview. Possible values: 1 (=yes), 2 (=no).

Conversion. If hefla==1 { Prospective falls (yes/no) = 1}, otherwise { Prospective falls (yes/no) = 0}.

Predictor target variables

Age

Source variable: "indager".

Description of "indager". Label in the dataset: 'Definitive age variable collapsed at 90 plus'. Possible values: non-negative integers less than 90.

Conversion. Age = indager.

Sex

Source variable: "indsex".

Description of "indsex". Label in the dataset: 'Definitive sex variable'. Possible values: 1 (=male), 2 (=female).

Conversion. Sex = indsex – 1.

Living situation

Source variable: "DhR", "DhR[2-12]"

Description of “DhR” and “DhR[2-12]”. Relationship of the subject with person i in the household, i=1,…,12. Possible values: 1 (=husband/wife), 2 (=partner/cohabitee), …, 96 (=self).

Conversion. The number of source variables that are not left blank is used to determine the number of subjects in the same household. If there is only one subject for a given household, {Living situation = 1}, otherwise { Living situation = 0}.

History of falls (yes/no)

Source variable: “HeFla”.

Description of “HeFla”. Whether the subject has ever fallen down since the last interview. Possible values: 1 (=yes), 2 (=no).

Conversion. If HeFla==1 { History of falls (yes/no) = 1}, otherwise { History of falls (yes/no) = 0}.

History of falls (number)

Source variable: “HeFla”, “HeFlb”.

Description of “HeFla” is given above.

Description of “HeFlb”. Number of falls experienced since the last interview.

Conversion. If HeFla == 2 { History of falls (number) = 0}, otherwise { History of falls (number) = HeFlb}.

Physical activity limitation

Source variable: “palevel”.

Description of “palevel”. Derived variable. Label in the dataset: ‘Physical activity summary’. Possible values: 0 (=Sedentary), 1 (=Low), 2 (=Moderate), 3 (=High). Description of variable from the ELSA documentation: “This variable summarises the answers to the level of work activity (WPJACT) in the work and pensions section and three questions on physical activity “in daily life” in the health section (HEACTA - HEACTC). It approximates as closely as possible the classification used in the Allied Dunbar Survey of Fitness. [Reference Activity and Health Research (1992) Allied Dunbar National Fitness Survey: main findings, London: Sports Council and Health Education Authority.] Levels used in this variable are defined as follows

- Sedentary: Not working or sedentary occupation, engages in mild exercise 1–3 times a month or less, with no moderate or vigorous activity.
- Low: Standing occupation, engages in moderate leisure-time exercise once a week or less and no vigorous activity; OR engages in mild leisure-time activity at least 1–3 times a month, moderate once a week or less and no vigorous; OR has a sedentary or no occupation and engages in moderate leisure-time activity once a week or 1–3 times a month, with no vigorous activity.
- Moderate: Does physical work; OR engages in moderate leisure-time activity more than once a week; OR engages in vigorous activity once a week to 1–3 times a month.
- High: Heavy manual work or vigorous leisure activity more than once a week.”

For knowing the exact algorithm, please refer to the ELSA documentation.

Conversion. If palevel == 0 { Physical activity limitation = 1}, otherwise { Physical activity limitation =0}.

Physical disability

Source variables: “headb[01-13]”.

Description of “headb[i]”, i=01,02,…,13. Label in the dataset ‘IADL: activity has problem with due to health/physical problem (i-th mention)’. Possible values:

- 1 = Dressing, including putting on shoes and socks
- 2 = Walking across a room
- 3 = Bathing or showering
- 4 = Eating, such as cutting up your food
- 5 = Getting in or out of bed
- 6 = Using the toilet, including getting up or down
- 7 = Using a map to get around in a strange place
- 8 = Preparing a hot meal
- 9 = Shopping for groceries
- 10 = Making telephone calls
- 11 = Taking medications
- 12 = Doing work around the house or garden
- 13 = Managing money, eg paying bills & keeping track ofexpenses
- 96 = None of these

Conversion. If at least a problem is recorded on any of the activities marked by values 2-6 { Physical disability = 1}, otherwise { Physical disability = 0}.

Instrumental disability

Source variable: “headb[01-13]”.

Description of these variables is given above.

Conversion. If at least a problem is recorded on any of the activities marked by values 8-13 { Physical disability = 1}, otherwise { Physical disability = 0}.

Cognitive impairment

Source variable: “hedibde”.

Description of “hedibde”. This variable shows whether a respondent has ever reported dementia (senility or another serious memory impairment) and if so, by which wave. Label in the dataset: ‘Ever reported dementia or memory impairment (diagnosed)’. Derived from question like ‘Has a doctor told you (or name of the subject) that you have any of the conditions in this card? …Dementia, senility or another serious memory impairment’. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Conversion. If hedibde == 1 or 2 { Cognitive impairment =1}, otherwise { Cognitive impairment =0}.

Note. Another source could be “cfind”. Label of “cfind”: ‘Total Cognitive Index (Memory + Executive)’. Brief description: ‘This variable gives the total cognitive score, which is the sum of the Memory and Executive indices’. Maybe this would be more appropriate.

Depression

Source variable: “totpsc”.

Description of “totpsc”. Total score on 8-item CESD scale. Label in the dataset ‘Eligibility for psfeel (Sum of all eight CES-D items, psceda - pscedh)’. Possible values: 0, 1, 2,…, 8.

Conversion. If totpsc >= 5 {Depression = 1}, otherwise {Depression = 0}. Threshold taken from the questionnaire, 5050_Wave_2_Documentation.pdf page 416/953.

History of stroke

Source variable: “hedimst”.

Description of “hedimst”. This variable shows whether a respondent has ever reported a stroke (cerebral vascular disease) and if so, by which wave. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Conversion. If hedimst == 1 or 2 { History of stroke =1}, otherwise { History of stroke =0}.

Urinary incontinence

Source variable: “HeInct”.

Description of “HeInct”. Question: ‘We would like to ask you about incontinence. During the last 12 months, have you lost any amount of urine beyond your control?’. Possible values: 1 (=yes), 2 (=no).

Conversion. If HeInct == 1 { Urinary incontinence =1}, otherwise { Urinary incontinence =0}.

Rheumatic disease

Source variable: “hedibar”.

Description of “hedibar”. This variable shows whether a respondent has ever reported arthritis (including osteoarthritis and rheumatism) and if so, by which wave. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Conversion. If hedibar == 1 or 2 { Rheumatic disease =1}, otherwise { Rheumatic disease =0}.

Dizziness and vertigo

Source variables: “HeFunc”, “HeAtt[01-14]”, “HeDiz”.

Description of “HeFunc”. Question: ‘By yourself and without using any special equipment, how much difficulty do you have walking for a quarter of a mile?’ Possible values: 1 (=no difficulty), 2(=some difficulty), 3 (=much difficulty), 4 (=unable to do this).

Description of “HeAtt[i]”, i=01, 02,…, 14. Label: ‘Symptoms that make walking 1/4 mile difficult (i-th mention)’. Possible values:

- 1 = Chest pain
- 2 = Fatigue/too tired
- 3 = Shortness of breath
- 4 = Tremor(s)
- 5 = Pain in leg or foot
- 6 = Swelling in leg or foot
- 7 = Incontinence or fear of incontinence
- 8 = Seeing difficulty

- 9 = Hearing difficulty
- 10 = Confusion
- 11 = Difficulty concentrating
- 12 = Memory problems
- 13 = Unsteady on feet or balance problems
- 14 = Lightheaded or dizziness
- 15 = Fear of falling
- 16 = Anxiety or fear
- 17 = Amputation
- 95 = Some other problem or symptom

Description of "HeDiz". Question: 'How often do you have problems with dizziness when you are walking on a level surface?'. Possible values: 1 (= Always), 2 (= Very often), 3 (= Often) 4 (= Sometimes), 5 (= Never) 6 (= SPONTANEOUS: Never walks), 7 (= SPONTANEOUS: Can't walk).

Conversion.

difficulty = (HeFunc == 2 or 3 or 4).

If (difficulty is true and any of the HeAtt[i] == 13 or 14) or (HeDiz == 1 or 2 or 3 or 4) { Dizziness and vertigo =1}, otherwise { Dizziness and vertigo =0}.

Diabetes

Source variable: "hedimdi".

Description of "hedimdi". Label: 'Ever reported diabetes or high blood sugar (diagnosed)'. This variable shows whether a respondent has ever reported diabetes or high blood sugar and if so, by which wave. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Conversion. If hedimdi == 1 or 2 { Diabetes =1}, otherwise { Diabetes =0}.

Note. Other possible source variables: "hedbts" ('Ever reported diabetes (diagnosed)'), "fglu" (Blood glucose level (mmol/L) - fasting samples only), "hba1c" ('Blood glycated haemoglobin level (%)').

Self-perceived health status

Source variable: none found.

Harmonization not possible.

Pain

Source variables: "HeYRa", "claud", "HePain".

Description of "HeYRa". Label: 'Whether had angina or chest pains in last 2 years'. Possible values: 1 (=yes), 2 (=no).

Description of "claud". This variable gives the summary value for the Edinburgh claudication scale, which relies on report of symptoms. To qualify as having claudication a respondent has to i) experience pain or discomfort on walking ii) not get it when standing still or sitting iv) the pain disappears in 10 minutes or less when stop v) experience the pain in the calf. Grade 1 applies if experiences the pain when walking uphill or

when in a hurry; grade 2 applies if experiences the pain when walking at an ordinary pace on level ground. Possible values: 0 (=none), 1 (=grade 1), 2 (=grade 2).

Description of “HePain”. Question: ‘Are you often troubled with pain?’. Possible values: 1 (=yes), 2 (=no).

Conversion.

Chest_pain = (HeYRa == 1)

Claudication = (claud == 1 or 2)

Generic_pain = (HePain == 1)

If Chest_pain or Claudication or Generic_pain is true {Pain = 1}, otherwise {Pain = 0}.

Note. Chest pain is excluded when constructing Pain in InCHIANTI.

Fear of falling

Source variables: “HeFunc”, “HeAtt[01-14]”.

Descriptions of “HeFunc” and “HeAtt[01-14]” are given in section about Dizziness and vertigo.

Conversion.

difficulty = (HeFunc == 2 or 3 or 4).

If (difficulty is true and any of the HeAtt[i] == 15) { Fear of falling =1}, otherwise { Fear of falling =0}.

Parkinson disease

Source variable: “hedibpd”.

Description of “hedibpd”. This variable shows whether a respondent has ever reported Parkinson’s disease and if so, by which wave. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Conversion. If hedibpd == 1 or 2 { Parkinson disease =1}, otherwise { Parkinson disease =0}.

Number of medications

Source variables: “Hemda”, “HeBetb”, “HeMdb”, “HeIns”, “HeAcea”, “Hehrtb”, “HeLng”, “HeAma”, “HeOstea”, “HeOstec”, “HePsya”, “HePsyb”, “HePad”, “Hehrtb2”, “Henictk”, “PsPsya”, “PsPsyb”.

Description of “Hemda”. Label: ‘High blood pressure: whether taking medication’. Possible values: 1 (yes), 2 (=no).

Description of “HeBetb”. Label: ‘Myocardial infarction: whether taking beta-blocker (coded by interviewer)’. Possible values: 1 (=taking beta-blockers), 2 (=not taking beta-blockers), 3 (=Taking other beta-blocker not on the showcard).

Description of “HeMdb”. Label: ‘Diabetes: whether taking medication’. Possible values: 1 (=yes), 2 (=no).

Description of “HeIns”. Label: ‘Diabetes: whether currently injects insulin’. Possible values: 1 (=yes), 2 (=no).

Description of “HeAcea”. Label: ‘Diabetes: whether taking ACE inhibitor or A2 receptor blocker’. Possible values: 1 (=Taking ACE inhibitor or A2 receptor blocker), 2 (=not taking ACE inhibitor or A2 receptor blocker), 3 (Taking ACE inhibitor not on the showcard).

Description of “Hehrtb”. Label: ‘Angina or MI: whether taking anticoagulant medication’. Possible values: 1 (=yes), 2 (=no).

Description of “Hehrtb2”. Label: ‘Angina-type pain: whether taking anticoagulant medication’. Possible values: 1 (=yes), 2 (=no).

Description of “HeLng”. Label: ‘Whether taking medication for lung condition’. Possible values: 1 (=yes), 2 (=no).

Description of “HeAma”. Label: ‘Whether taking medication for asthma’. Possible values: 1 (=yes), 2 (=no).

Description of “HeOstea”. Label: ‘Osteoporosis: whether takes calcium pills or vitamin D’. Possible values: 1 (=yes), 2 (=no).

Description of “HeOstec”. Label: ‘Osteoporosis: whether taking medication’. Possible values: 1 (=yes), 2 (=no).

Description of “Henictk”. Label: ‘Smoking: whether taking medication to help stop’. Possible values: 1 (=yes), 2 (=no).

Description of “HePad”. Label: ‘Hip or knee pain: whether taking medication or doing exercise’. Possible values: 1 (=yes), 2 (=no).

Description of “HePsya”. Label: ‘Depression: action recommended by doctor (medication and/or counselling)’. Possible values: 1 (=medication), 2 (=counseling), 3 (=both medication and counseling), 4 (=none).

Description of “HePsyb”. Label: ‘Depression: whether started treatment within 2 weeks of it being offered’. Possible values: 1 (=yes), 2 (=no).

Description of “PsPsya”. Label: ‘Whether doctor/nurse suggested takes medication or sees specialist’. Possible values: 1 (=medication), 2 (=counseling), 3 (=both medication and counseling), 4 (=none).

Description of “PsPsyb”. Label: ‘Whether respondent started treatment within 2 weeks of being offered it’. Possible values: 1 (=yes), 2 (=no).

Conversion.

If Hemda==1 {medpress=1}, otherwise {medpress=0}. Similarly, meddiab, medins, medantic1, medantic2, medlung, medasth, medost1, medost2, medhkp, medsmok are constructed respectively from HeMdb, HeIns, Hehrtb, Hehrtb2, HeLng, HeAma, HeOstea, HeOstec, HePad, Henictk.

If HeBetb==1 or 3 {medbblock=1}, otherwise { medbblock=0}. Similarly meddiabACE is constructed from HeAcea.

If (HePsya==1 or 3) and (HePsyb==1) {meddep1 = 1}, otherwise {meddep1 = 0}. Similarly, meddep2 is constructed from PsPsya and PsPsyb.

Number of medications is constructed as the sum of medpress, meddiab, medins, medantic1, medantic2, medlung, medasth, medost1, medost2, medhkp, medsmok, medbblock, meddiabACE, meddep1, meddep2.

Note. The construction of this variable is particularly complex. It is likely that it underestimates the real number of medications because subjects are not asked on all the drug classes. Furthermore, on the principal component analysis led on the three datasets stacked together, "Number of medications" has the highest loading on the second principal component. As the second principal component well separates the ELSA population from the other two, this is likely to be a sign that the construction of the variable "Number of medication" was not successful on the ELSA dataset. One possibility could be just to drop this target variable from ELSA and say that in ELSA harmonization on this variable is not believed possible.

Sedatives

Source variable: none found.

Harmonization not possible.

Antihypertensives

Source variable: none found.

Harmonization not possible.

Antiepileptics

Source variable: none found.

Harmonization not possible.

Grip strength

Source variables: "mmgsd1", "mmgsn1", "mmgsd2", "mmgsn2", "mmgsd3", "mmgsn3".

Description of "mmgsd1", "mmgsn1", "mmgsd2", "mmgsn2", "mmgsd3", "mmgsn3". Isometric handgrip strength measure, obtained with a grip gauge. Three measures are obtained from the dominant hand ("mmgsd[1-3]"), and three from the non-dominant hand ("mmgsn[1-3]"). Unit of measurement: Kg.

Conversion. Grip strength is computed as the maximum between the mean grip strength measured on the dominant hand, and the mean grip strength measured on the non-dominant hand.

Gait speed

Source variables: "MMWlkA", "MMWlkB".

Description of "MMWlkA", "MMWlkB". Time to complete 8 feet (= 2.4384 m) walking test at self-selected gait speed, first and second trial.

Conversion.

speed_8f = 2 * 2.4384/ (MMWlkA + MMWlkB)

If speed_8f < 1 {Gait speed = 0.01 + 1.052 * speed_8f}, otherwise { Gait speed = 0.481 + 0.581 * speed_8f}.

Note. This formula, used to compare gait speed obtained from 8-feet test to one obtained from a 4-m test, was proposed in Guralnik et al 2000, 'Lower Extremity Function and Subsequent Disability : Consistency

Across Studies , Predictive Models , and Value of Gait Speed Alone Compared With the Short Physical Performance Battery' and was further used in Studenski et al 2011, 'Gait speed and survival in older adults.'

SPPB balance

Source variables: "mmssre", "mmstre", "mmftti", "mmftre2".

Description of "mmssre". Label: 'Side-by-side stand: Outcome'. Possible values: 1 (=Held for 10 seconds), 2 (= Held for less than 10 seconds), 3 (= Not attempted).

Description of "mmstre". Label: 'Semi-tandem stand: Outcome'. Possible values: 1 (=Held for 10 seconds), 2 (= Held for less than 10 seconds), 3 (= Not attempted).

Description of "mmftti". Label: 'Full tandem stand: Time position held (seconds)'.

Description of "mmftre2". Label: '(D) Outcome of full tandem stand according to age'. Possible values: -1 (=Ineligible - did not hold semi-tandem stand for 10 seconds), 1 (=Held for 10 seconds, respondent aged 70 or over), 2 (=Held for less than 10 seconds, respondent aged 70 or over), 3 (=Held for 30 seconds, respondent aged less than 70), 4 (=Held for less than 30 seconds, respondent aged less than 70), 5 (=Stand not attempted).

Conversion.

If side-by-side not attempted or held for less than 10 s (i.e. if mmssre== 2 or 3).{SPPB balance = 0}.

If side-by-side held for 10 s but semi-tandem not attempted or held for less than 10 s (i.e. if mmssre==1 and mmstre == 2 or 3) {SPPB balance = 1}.

If side-by-side and semi-tandem successfully held for 10 s, but tandem test not attempted or held for less than 3 s {SPPB balance = 2}.

If side-by-side and semi-tandem successfully held for 10 s, and tandem held for at least 3 s but less than 10 s {SPPB balance = 3}.

If side-by-side, semi-tandem, and tandem successfully held for 10 s {SPPB balance = 4}.

SPPB gait

Source variables: "MmTrya", "MmTryb", "MMWlkA", "MMWlkB".

Description of "MmTrya", "MmTryb". Outcome of first and second walk respectively. Possible values: 1 (= Completed successfully), 2 (= Attempted but unable to complete), 3 (= Stopped by the interviewer because of safety reasons), 4 (= Respondent refused).

Description of "MMWlkA", "MMWlkB" is already given in section about Gait speed.

Conversion.

maxspeed_8f = 2.4384/ min(MMWlkA, MMWlkB)

If maxspeed_8f < 1 {maxspeed = 0.01 + 1.052 * maxspeed_8f}, otherwise { maxspeed = 0.481 + 0.581 * maxspeed_8f}. [see section on Gait speed for more clarifications].

If MmTrya or MmTryb == 2 or 3 {SPPB gait = 0}.

If maxspeed <= 0.43 {SPPB gait = 1}.

If maxspeed > 0.43 and maxspeed <= 0.60 {SPPB gait = 2}.

If maxspeed >0.60 and maxspeed <= 0.77 {SPPB gait = 3}.

If maxspeed > 0.77 {SPPB gait = 4}.

SPPB chair standing

Source variables: “mmrrre”, “mmcrav”, “mmcrna”, “mmrrna”, “mmrrfti”.

Description of “mmrrre”. Label: ‘Chair rise: Outcome of multiple chair rises (number of rises completed)’.

Description of “mmcrav”. Label: ‘Chair rise: Whether suitable chair available’.

Description of “mmcrna”. Label: ‘Chair rise: Reason single chair rise not attempted’.

Description of “mmrrna”. Label: ‘Chair rise: Reason multiple chair rises not attempted’.

Description of “mmrrfti”. Label: ‘Chair rise: Time to complete 5 rises (seconds)’.

Conversion.

Variables “mmcrav”, “mmcrna”, “mmrrna” are used to determine when SPPB chair standing is not available.

If mmrrre < 5 {SPPB chair standing = 0}.

If mmrrfti >= 16.7 { SPPB chair standing = 1}.

If mmrrfti >= 13.6 and mmrrfti < 16.7 { SPPB chair standing = 2}.

If mmrrfti >= 11.2 and mmrrfti < 13.6 { SPPB chair standing = 3}.

If mmrrfti < 11.2 { SPPB chair standing = 4}.

SPPB

Source variable: harmonized variables “SPPB balance”, “SPPB gait”, “SPPB chair standing”.

Description of the harmonized variables “SPPB balance”, “SPPB gait”, “SPPB chair standing” is given above.

Conversion. SPPB = SPPB balance + SPPB gait + SPPB chair standing.

Gait problems

Source variable: harmonized variable “SPPB gait”.

Description of “SPPB gait” is given above.

Conversion. If SPPB gait < 4 {Gait problems = 1}, otherwise {Gait problems = 0}.

Walking aid use

Source variables: “HeAid[1:5, 9:12, 17:22]”, “MmAid”.

Description of “HeAid[1:5, 9:12, 17:22]”. Label: ‘Walking aids used’, different mentions and in different parts of the questionnaire (disability, balance, ADL-IADL). Possible values: 96 (= None of these), 1(= A cane or walking stick), 2 (= A zimmer frame or walker), 3 (= A manual wheelchair), 4 (= An electric wheelchair), 5 (= A buggy or scooter), 6 (= Special eating utensils), 7 (= A personal alarm), 8 (= Elbow crutches).

Description of “MmAid”. Type of aid used during the walking test. 1.0 (= None), 2 (= Walking stick or cane), 3 (= Elbow crutches), 4 (= Walking frame), 5 (= Other), 85 (= Other answer - not codeable 01 to 04), 86 (= Irrelevant response - not codeable 01 to 04).

Conversion. If (any of the HeAid[1:5, 9-12, 17:22] == 1, 2, 3, 4, 5 or 8) or (MmAid == 2, 3, 4, 5 or 85) {Walking aid use = 1}, otherwise {Walking aid use = 0}.

Vision impairment

Source variable: “Heeye”, “heoptgl”, “heoptdi”, “heoptmd”.

Description of “Heeye”. Label: ‘Self-reported eyesight (while using lenses, if appropriate)’. Possible values: 1 (= excellent), 2 (= very good), 3 (= good), 4 (= fair), 5 (=poor), 6 (= SPONTANEOUS: Registered or legally blind).

Description of “heoptgl”. This variable shows whether a respondent has ever reported glaucoma and if so, by which wave. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Description of “heoptdi”. This variable shows whether a respondent has ever reported diabetic eye disease and if so, by which wave. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Description of “heoptmd”. This variable shows whether a respondent has ever reported macular degeneration and if so, by which wave. Possible values: 0 (=neither waves), 1 (=by wave 1), 2 (=by wave 2).

Conversion.

If Heeye == 4 or 5 or 6 {imp1 = 1}, otherwise {imp1 = 0}.

If heoptgl == 1 or 2 {glaucoma = 1}, otherwise {glaucoma = 0}. Similarly, diabeticed and maculardeg are constructed from heoptdi and heoptmd.

If imp1 or glaucoma or diabeticed or maculardeg == 1, {Vision impairment =1}, otherwise {Vision impairment = 0}.

Note. Similarly to Number of medication, Vision impairment has a great loading on the second principal component of the PCA run on the three datasets stacked together.

Other possible source variables:

- “Hefrnd” (Eyesight for recognition of friend across street)
- “Hepap” (Eyesight for reading ordinary newspaper print)
- “fqhelp” (Any help respondent needed with reading the showcards during the interview)
- “heoptca” ((D) Ever reported cataract (diagnosed)).

Hearing impairment

Source variable: “Hehear”.

Description of “Hehear”. Self-reported hearing (while using hearing aid if appropriate). Possible values: 1 (= excellent), 2 (= very good), 3 (= good), 4 (= fair), 5 (=poor).

Conversion. If Hehear == 4 or 5 {Hearing impairment = 1}, otherwise {Hearing impairment = 0}.

Appendix 3. InCHIANTI

This appendix contains a description of how the variables in the InCHIANTI dataset have been processed to validate FRAT-up.

The variables of the InCHIANTI dataset are called 'source variables'. The variables that are needed in FRAT-up are called 'target variables'. Among the target variables we distinguish the predictor target variables and an outcome target variable. In Chapter 4, predictor target variables are also called 'risk factors' and the source variables used to derive these target variables are called 'risk factor estimators'.

For the sake of clarity and brevity, the documentation does not take into account in details how missing values (not applicable, refusal, unknown…) are coded and handled. To have this kind of information, please refer to the original documentation of the database and contact us for the R scripts.

Outcome target variable

Prospective falls (yes/no)

Source variable: "IY15_V1"

Description of "IY15_V1". Question asked at follow-up 1: 'Did you ever fall down in the last 12 months?'. Follow-up 1 was led about 3 years after the baseline assessment. Possible values: 0 (=no), 1 (=yes).

Conversion. Prospective falls (yes/no) = IY15_V1

Predictor target variables

Age

Source variable: "IXAGE".

Description of "IXAGE". Age in years at interview. Possible values: non-negative integers.

Conversion. Age = IXAGE.

Sex

Source variable: "SEX"

Description of "SEX". Gender. Possible values: 1 (=male), 2 (=female).

Conversion. Sex = SEX-1.

Living situation

Source variables: "IX3_V[11+i 5]", "IX3_V[12+i 5]", i =0,1,2,…,11.

Description of "IX3_V[11+i 5]", i =0,1,2,…,11. Question: 'Indicate each person you live with or see regularly (Person i+1)'. Possible values: 1 (=spouse), 2 (=son), etc.

Description of "IX3_V[12+i 5]", i =0,1,2,…,11. Question: 'Person i+1: Is this person living with you or not?'. Possible values: 0 (=not co-habitant), 1 (=co-habitant).

Conversion. If at least one person of the ones indicated in IX3_V[11+i 5] is co-habitant, as indicated in IX3_V[12+i 5], then Living situation = 1, otherwise, Living situation = 0.

History of falls (yes/no)

Source variable: “IX15_V1”.

Description of “IX15_V1”. Question: ‘Did you ever fall down in the last 12 months?’. Possible values: 0 (=no), 1 (=yes).

Conversion. History of falls (yes/no) = IX15_V1.

History of falls (number)

Source variables: “IX15_V1”, “IX15_V3”.

Description of “IX15_V1” is given above.

Description of “IX15_V3”. Question: ‘How many times did you fall down in the last 12 months?’. Possible values: 1,2,…,9.

Conversion. If IX15_V1==0 { History of falls (number) = 0}, otherwise { History of falls (number) = IX15_V3}.

Physical activity limitation

Source variable: “IX14_V26”.

Description of “IX14_V26”. Question: ‘Physical activity level last year’. Possible values:: 1=almost no activity (driven in bed or almost); 2=sitting for most of the time, rarely a short walk or other non-demanding activity; 3=low-intensity exercise (walking, dancing, hunting or fishing, do the shopping without car) at least 2-4 hours a week; 4=moderate-intensity exercise (running, walking uphill, swimming, gymnastic, hoeing in the garden, riding a bike uphill, etc.) for at least 1-2 hours a week or low-intensity exercise (see 3) for more than 4 h/week; 5=moderate-intensity exercise for more than 3h/week; 6=intense physical exercise, regularly, many times a week; 7=use this code for people that have walk much (at least 5Km/day), regularly (at least 5 days/week) and for long (at least for 5 consecutive years).

Light-intensity exercise is defined as the exercise that does not come with sweating and can also be practiced while talking with another person; moderate-intensity exercise as the exercise that is associated with sweating and does not allow talking at the same time; intense exercise is the maximal one, to the limit of endurance.

Conversion. If IX14_V26 <= 2 { Physical activity limitation = 1}, otherwise { Physical activity limitation = 0}.

Physical disability

Source variable: “IXADL_T”.

Description of “IXADL_T”. Number of activity-of-daily-living (ADL) disabilities. Possible values: 0,1,2,…,6. Subjects are asked on the following activities:

- WHO activity 6: Any difficulty washing face and arms?
- WHO activity 19: Any difficulty controlling urination and bowel movements?
- WHO activity 8: Any difficulty dressing and undressing?
- WHO activity 12: Any difficulty getting in and out of bed?

- WHO activity 9: Any difficulty eating (e.g., holding a fork, cutting food, drinking from a glass)?
- WHO activity 11: Any difficulty using the toilet?

Conversion. If IXADL_T > 0 { Physical disability = 1}, otherwise { Physical disability = 0}.

Instrumental disability

Source variable: “IXIADL_T”.

Description of “IXIADL_T”. Number of instrumental-activity-of-daily-living (IADL) disabilities. Possible values: 0,1,2,…,8. Subjects are asked on the following activities:

- WHO activity 20: Any difficulty using the telephone?
- WHO activity 22: Any difficulty using public transportation?
- WHO activity 10: Any difficulty cooking a simple meal?
- WHO activity 13: Any difficulty doing light housework (e.g., doing dishes, light cleaning)?
- WHO activity 14: Any difficulty doing heavy housework (e.g., washing windows, floor)?
- WHO activity 23: Any difficulty taking medications correctly?
- WHO activity 24: Any difficulty managing home finances?
- WHO activity 5: Any difficulty shopping daily for basic necessities?

Conversion. If IXIADL_T > 0 { Instrumental disability = 1}, otherwise { Instrumental disability = 0}.

Cognitive impairment

Source variable: “IXMMSECR”.

Description of “IXMMSECR”. Mini-mental state examination (MMSE) raw score. Possible values: =0, 1, 2, …, 30.

Conversion. If IXMMSECR <= 20 {Cognitive impairment = 1}, otherwise { Cognitive impairment = 0}.

Depression

Source variable: “IXCESD_T”.

Description of “IXCESD_T”. Total score on 20-item CESD scale. Possible values: 0, 1, 2, …, 60.

Conversion. If IXCESD_T > 21 { Depression = 1}, otherwise { Depression = 0}.

History of stroke

Source variable: “AXSTROKE”.

Description of “AXSTROKE”. Ascertained history of stroke. Possible values: 0 (=no evidence), 1 (=definite), 2 (=possible), 3 (=TIA).

Conversion. If AXSTROKE == 1 { History of stroke = 1}, if AXSTROKE == 2 { History of stroke = not available}, if AXSTROKE == 0 or 3 { History of stroke = 0}.

Urinary incontinence

Source variable: “IX13_V6”.

Description of “IX13_V6”. Question: ‘Over the last year did you ever lose control of urine? (for example, while coughing)’. Possible values: 0 (=no), 1 (=yes).

Conversion. Urinary incontinence = IX13_V6.

Rheumatic disease

Source variables: “AXGONART”, “AXANCART”.

Description of “AXGONART”. Knee arthritis. Possible values: 0 (=no evidence), 1 (=Knee replacement/ Pain + stiffness), 2 (=Pain OR stiffness, not both).

Description of “AXANCART”. Hip arthritis. Possible values: 0 (=no evidence), 1 (=Hip replacement/ Pain + stiffness), 2 (=Pain OR stiffness, not both).

Conversion. If ((AXGONART == 1 or 2) or (AXANCART == 1 or 2)) { Rheumatic disease = 1}, otherwise { Rheumatic disease = 0}.

Dizziness and vertigo

Source variable: “VX10_V22”.

Description of “VX10_V22”. Question ‘Have you ever experienced dizziness or unsteadiness in last year?’. Possible values: 0 (=no), 1 (=yes).

Conversion. Dizziness and vertigo = VX10_V22.

Diabetes

Source variable: “AXDIAB2A”.

Description of “AXDIAB2A”. Diabetes mellitus (incl. bl glucose>=126). Possible values: 0 (=no evidence), 1 (=definite), 2 (=possible).

Conversion. If AXDIAB2A == 0 { Diabetes = 0}, otherwise { Diabetes = 1}.

Self-perceived health status

Source variable: “IX8_V1”.

Description of “IX8_V1”. Question: ‘How would you evaluate your current health? How do you feel now?’. Possible values: 1 (=very poor), 2 (=poor), 3 (=fair, so and so), 4 (=good), 5 (=very good).

Conversion. If IX18_V1 <= 2 { Self-perceived health status = 1}, otherwise { Self-perceived health status = 0}.

Pain

Possible source variables: “VX8_V26”, “VX11_V9”, “VX12_V6”, “VX16_V1”, “VX17_V3”, “IX12_V15”, “IX19_V11”, “IX19_V12”.

Description of “VX8_V26”. Question: ‘Stomach pain in last year...how severe?’. Possible values: 0 (=No problem), 1 (=Light, only aware if think about it), 2 (=Moderate, does not interfere w/ life), 3 (=Interferes with daily living), 4 (=Severe, unable to have a normal life).

Description of “VX11_V9”. Chest pain resting/following exertion. Possible values: 0 (=no), 1 (=yes).

Description of “VX12_V6”. Question ‘Ever had pain in the legs while walking?’. Possible values: 0 (=no), 1 (=yes, currently), 2 (=yes, in the past).

Description of “VX16_V1”. Question ‘Ever had back pain in last 12 months?’ Possible values: 0 (=no), 1 (=yes).

Description of “VX17_V3”. Question ‘During the last 5 years, have you had pain in knees/hips requiring medications?’. Possible values: 0 (=no), 1 (=yes).

Description of “IX12_V15”. Question ‘How often in the last month did you have pain while sleeping (e.g., muscular cramps)?’. Possible values: 1 (=Never during the past month), 2 (=Less than once a week), 3 (=Once or twice a week), 4 (=Three or more times a week).

Description of “IX19_V11” (“IX19_V12”). Question ‘At present, do you ever experience pain in the right (left) foot?’. Possible values: 1 (=Never), 2 (=Sometimes), 3 (=Often), 4 (=Always).

Conversion.

legpain = VX12_V6==1

backpain = VX16_V1==1

kneehippain = VX17_V3==1

sleeppain = IX12_V15>=2

rfootpain = IX19_V11>=2

lfootpain = IX19_V12>=2

If legpain or backpain or kneehippain or sleeppain or rfootpain or lfootpain is true {Pain = 1}, otherwise {Pain = 0}.

Note. “VX8_V26” and “VX11_V9” not used.

Fear of falling

Source variables: “IX16_V[1+i 4]”, “IX16_V[2+i 4]”, “IX16_V[4+i 4]”, for i=0,1,2,…,10.

Description of “IX16_V[1+i 4]”. Question ‘Do you usually do activity i?’. Possible values: 0 (=no), 1 (=yes).

Description of “IX16_V[2+i 4]”. Question ‘If you usually do activity i, are you afraid of falling while doing it?’. Possible values: 0 (=no), 1 (=yes).

Description of “IX16_V[4+i 4]”. Question ‘If you do not do activity i, don’t you do it because you are afraid of falling?’. Possible values: 0 (=no), 1 (=yes).

The activities on which subjects are asked are:

- going out to shop
- cooking their meals
- bathing or showering without help
- getting out of the bed without help
- taking a walk outside the house without help

- walking on slippery surfaces without help
- going to visit relatives or friends without help
- reaching up by themselves for something located high over their head
- going to crowded places by themselves
- taking long walks (1 hour or more) by themselves (without help)
- bending down to pick up an object from the floor

The whole questionnaire is one version of the SAFE (Survey of Activities and Fear of Falling in the Elderly).

Conversion. If there is at least one activity that is not performed because of fear of falling { Fear of falling = 1}, otherwise { Fear of falling = 0}.

Parkinson disease

Source variable: “AXPARK”.

Description of “AXPARK”. Parkinson's disease. Possible values: 0 (=no evidence), 1 (=definite), 2 (=possible).

Conversion. If AXPARK == 0 { Parkinson disease = 0}, if AXPARK == 1 { Parkinson disease = 1}, if AXPARK == 2 { Parkinson disease = not available}.

Number of medications

Source variable: “IXN_FARM”.

Description of “IXN_FARM”. Number of medications. Possible values: 0, 1, 2, …, 10.

Conversion. Number of medications = IXN_FARM.

Sedatives

Source variable: “IXATC_[i]”, for i=1,2,…,10..

Description of “IXATC_[i]”. ATC code of the medications used by the subject.

Conversion. If for any i=1,2,…,10 IXATC_[i] starts with “N05C” {Sedatives = 1}, otherwise {Sedatives = 0}.

Antihypertensives

Source variable: “IXATC_[i]”, for i=1,2,…,10..

Description of “IXATC_[i]” is given above.

Conversion. If at least one of the ATC codes is in the list (“C09CA[01-04,06-08]”, “C09DA[01-04,06,08”, “C03AA03”, “C03BA04”, “C03BA11”, “C03CA[01,03,04]”, “C03EA01”, “C03EB01”, “QC03CA0”, “C09AA”, “C09AA[01-13,15]”, “C09BA[01-09,12]”, “C08CA[01-09,11-13]”, “C08DA01”, “C08DB01”, “C07AA05”, “C07AA07”, “C07AA12”, “C07AB[02,03,07,08,12]”, “C07AG02”, “C07CA02”, “C07CB02”, “C07CB03”, “C02AC01”, “C02CA04”, “C02CA05”, “C02CA49”, “C02KA49”, “C02KD01”, “C02LA51”, “C03DA[01-03]” { Antihypertensives = 1}, otherwise { Antihypertensives = 0}.

Antiepileptics

Source variable: “FX1_N3”.

Description of “FX1_N3”. Use of antiepileptics. Possible values: 0 (=no), 1 (=yes). Constructed from “IXATC_[i]” (described above) with comparison with the list of ATC codes (“N03AA02”, “N03AA03”,

“N03AB02”, “N03AB52”, “N03AF01”, “N03AF02”, “N03AG01”, “N03AX12”, “N03AX16”). If at least one of the “IXATC_[i]” is in the above list, then “FX1_N3” is set to 1.

Conversion. Antiepileptics = FX1_N3.

Grip strength

Source variables: “PX10_V[37-40]”.

Description of “PX10_V[37-40]”. Hand grip strength measured with a dynamometer. Unit of measurement: Kg. Hand grip is measured twice on the left hand (“PX10_V[38, 40]”) and twice on the right hand (“PX10_V[37, 39]”).

Conversion. Grip strength is computed as the maximum between the mean grip strength measured on the right hand, and the mean grip strength measured on the left hand

Gait speed

Source variables: “PXWLK1A”, “PXWLK1B”.

Description of “PXWLK1A” and “PXWLK1B”. Time for walking 4 m at usual pace, measured on two trials. Unit of measurement: s.

Conversion. Gait speed = 4 * 2 / (PXWLK1A + PXWLK1B).

Note. Conversion rule taken for consistency with ActiFE. It can be changed in (4/PXWLK1A + 4/PXWLK1B)/2.

SPPB balance

Source variable: “PXSPSB”.

Description of “PXSPSB”. Label in the dataset: ‘EPESE Performance Balance Sub-score (0 - 4)’. Possible values: 0,1,2,3,4.

Conversion. SPPB balance = PXSPSB.

SPPB gait

Source variable: “PXSPSW”.

Description of “PXSPSW”. Label in the dataset: ‘EPESE Performance Walking Sub-score (0 - 4)’. Possible values: 0,1,2,3,4.

Conversion. SPPB gait = PXSPSW.

SPPB chair standing

Source variable: “PXSPSC”.

Description of “PXSPSC”. Label in the dataset: ‘EPESE Performance Repeated Chair Stands Sub-score (0 - 4)’. Possible values: 0,1,2,3,4.

Conversion. SPPB chair standing = PXSPSC.

SPPB

Source variable: “PXSPS”.

Description of “PXSPS”. Label in the dataset: ‘EPESE Summary Performance Score (0 - 12)’. Possible values: 0, 1, 2, …, 12.

Conversion. SPPB = PXSPS.

Gait problems

Source variable: “PXSPSW”.

Description of “PXSPSW” already given above.

Conversion. If PXSPSW < 4 {Gait problems = 1}, otherwise {Gait problems = 0}.

Walking aid use

Source variable: “PX1_V10”.

Description of “PX1_V10”. Whether the subject walks independently (without aids, orthoses, prostheses), assessed during the physical therapy visit. Possible values: 0 (=no), 1 (=yes).

Conversion. If PX1_V10 == 0 {Walking aid use = 1}, otherwise {Walking aid use = 0}.

Vision impairment

Source variables: “VX24_V[36-46]”

Description of “VX24_V36”. Label in the dataset: ‘Visual acuity,3 meter(Monoyer 1/10-11/10)’. Possible values: 1(=1/10), 2 (=2/10), …11 (=11/10).

Description of “VX24_V37”. Label in the dataset: ‘Contrast sensitivity (0.05-2.0)’.

Description of “VX24_V[38-46]”. Nine tests of visual stereognosis. Possible values for each variable: 0 (= incorrect), 1 (= correct).

Conversion.

> acuity_impairment = VX24_V36 <= 5
>
> contrast_sensitivity_impairment = VX24_V37 <= 1.6
>
> stereo_total_score = sum(VX24_V[38-46])
>
> stereo_impairment = stereo_total_score <= 3
>
> If acuity_impairment or contrast_sensitivity_impairment or stereo_impairment is true { Vision impairment = 1}, otherwise { Vision impairment = 0}.

Hearing impairment

Source variables: “VX15_V1”.

Description of “VX15_V1”. Question ‘Do you have any trouble hearing?’. Possible values: 0 (=No), 1 (=Slight deafness), 2 (=Severe deafness), 3 (=Conversation impossible).

Conversion. If VX15_V1 >= 1 { Hearing impairment = 1}, otherwise { Hearing impairment = 0}.

Appendix 4

InCHIANTI SAS database

Table 1. List of all the tables of the SAS database. The extension of the files is *sas7bdat*.

Name of the table (number of variables)				Brief description
Ana_raw (76)				Vital status
Hosp2kxg (56)				Hospitalizations
Relatives (8)				Relatives
Baseline	**Follow-up 1**	**Follow-up 2**	**Follow-up 3**	
Int_rawe (892)	Inf1rawe (840)	Inf2rawe (862)	Inf3rawe (865)	Home interview
Cli_rawe (785)	Clf1rawe (761)	Clf2rawe (761)	Clf3rawe (761)	Clinical visit
Adju_ana (50)	Adjf1ana (31)	Adjf2ana (31)	Adjf3ana (31)	Diseases adjudicated via algorithms
Per_ana (100) Per_rawe (435)	Pef1_ana (93) Pef1rawe (401)	Pef2_ana (100) Pef2rawe (440)	Pef3_ana (100) Pef3rawe (471)	Physical Exam (Performance based tests of balance, gait, manual dexterity, ROM, muscle strength and muscle power)
Fmc_ana (88)	Fmcf_ana (88)	Fmcf2ana (88)	Fmcf3ana (90)	Drugs
Sup_raw (43)	Supf1raw (43)	Supf2raw (43)	Supf3raw (43)	Dietary supplements
Labo_raw (331) Oldtnfa (8)	Labf1raw (113) Othassay (7)	Labf2raw (119)	Labf3raw (81)	Laboratory assays
Mar_raw (74)	Marf1raw (96)	Marf2raw (102)	Marf3raw (102)	Instrumental exams (ECG, ENG, anthropometric measures, Eco-Color-Doppler, Blood pressure)
Pqct_raw (66)	Pqcf1raw (61)	Pqcf2rwn (61)	Pqcf3raw (61)	Peripheral quantitative computed tomography
		Biaf2rwn (42)	Biaf3raw (42)	Bioelectrical impedance analysis
Nutr_raw (48) Alim_raw (126) Epic_raw (486)	Nutf1raw (47) Alif1raw (124) Epif1raw (484)	Nutf2raw (47) Alif2raw (124) Epif2raw (484)	Nutf3raw (47) Alif3raw (124) Epif3raw (484)	Nutrition (EPIC questionnaire, nutrition habits, macro- and micro-nutrients intake)
Ped_rawe (410)				Podiatry assessment

Lasso model

Table 2. Variables that have been selected more frequently in the 10-fold validation procedure of the Lasso model, and their standardized regression coefficients.

	Number of times it was selected	Mean standardized regression coefficient
Number of previous falls	10	0.17
Number of drugs	10	0.08
Self-perceived health status	10	-0.07
Previous falls (yes/no)	10	0.06
Drugs for dementia (yes/no)	10	0.06
CESD depressed mood scale (0-28)	10	0.04
Q: "If you are retired, do you have a veteran pension?" (yes/no)	10	0.03
Q: "Can you walk 300 meters twice without stopping?" (yes/no)	10	-0.03
Gait speed, 4m usual pace	6	-0.03
anti-hypertensive (yes/no)	7	0.02
Q: "Do you have difficulty walking 400 meters on rough terrain?"	6	0.02
Psychoanaleptics: antidepressants (yes/no)	6	0.02
Walking posture: cautious attitude? (yes/no)	6	0.02
Family med hx: siblings diabetic? (yes/no)	7	0.02
Q: "Must you hold onto something (e.g., bannister) while climbing stairs?" (yes/no)	8	0.02
Quinolone antibacterials (yes/no)	6	0.02
Miscellaneous antihypertensives including alpha-blocking agents (yes/no)	6	$8 \cdot 10^{-6}$

Table 3. Marginal calibration assessment. Observed and predicted number of samples reporting a given number of falls. Error = observed – predicted. Relative error= (observed - predicted)/total number of samples.

Number of falls	0	1	2	3	4	5	6	7	8	9 or more	Total
Observed	1814	303	91	52	23	12	2	0	3	13	2313
Predicted	1780.0	328.6	112.3	46.0	21.0	10.4	5.6	3.2	1.9	4.1	
Error	34.0	-25.6	-21.3	6.0	1.2	1.6	-3.6	-3.2	1.1	8.9	
Relative error (x 100)	1.5	-1.1	-0.9	0.3	0.1	0.1	-0.2	-0.1	0.0	0.4	

Learning curves

A sensitivity analysis of the performance of the Lasso model to the number of samples available in the training dataset is carried out with repeated random sub-sampling validation. The data are split in a training

set and a test set. The split is repeated assigning different percentages of samples to the training test. For each percentage of samples included in the training test, the split is repeated randomly 162 times. Each time, the Lasso model is fitted in the training set and assessed in the test set. For each percentage of samples assigned to the training set, the performance of the model is taken as the mean of the performance achieved across the random splits.

All splits in folds, training and test sets are done so that all the samples relative to one same subjects are consistently assigned to the same set.

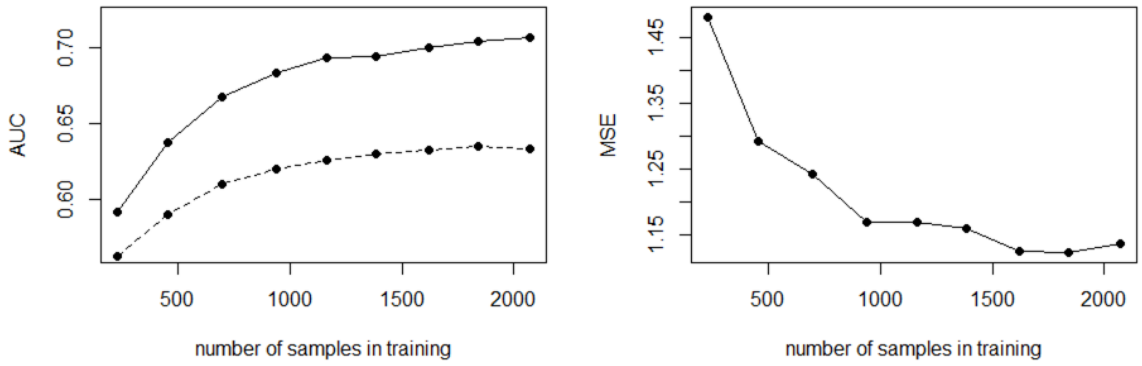

Figure 1. Learning curves. Sensibility of the performance of the trained models on the sample size of the training set. Left: AUC. Right: MSE.