Wearable Inertial Sensors for Fall Risk Assessment and Prediction in Older Adults: A Systematic Review and Meta-Analysis

Article *in* IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society · January 2018

DOI: 10.1109/TNSRE.2017.2771383

Luis Montesinos
Tecnológico de Monterrey, Campus Ciudad de México

Rossana Castaldo
SDN Istituto di Ricerca Diagnostica e Nucleare

Leandro Pecchia
The University of Warwick

Wearable Inertial Sensors for Fall Risk Assessment and Prediction in Older Adults: A Systematic Review and Meta-Analysis

Luis Montesinos, *Student Member, IEEE*, Rossana Castaldo, *Student Member, IEEE,* and Leandro Pecchia, *Member, IEEE*

***Abstract*— Wearable inertial sensors have been widely investigated for fall risk assessment and prediction in older adults. However, heterogeneity in published studies in terms of sensor location, task assessed and features extracted is high, making challenging evidence-based design of new studies and/or real-life applications. We conducted a systematic review and meta-analysis to appraise the best available evidence in the field. Namely, we applied established statistical methods for the analysis of categorical data to identify optimal combinations of sensor locations, tasks and feature categories. We also conducted a meta-analysis on sensor-based features to identify a set of significant features and their pivot values. The results demonstrated that with a walking test, the most effective feature to assess the risk of falling was the velocity with the sensor placed on the shins. Conversely, during quite standing, linear acceleration measured at the lower back was the most effective combination of feature-placement. Similarly, during the sit-to-stand and/or the stand-to-sit tests, linear acceleration measured at the lower back seems to be the most effective feature-placement combination. The meta-analysis demonstrated that four features resulted significantly higher in fallers: the root-mean-square acceleration in the mediolateral direction during quiet standing with eyes closed (Mean Difference (MD): 0.01 g; 95% Confidence Interval (CI95%): 0.006 to 0.014); the number of steps (MD: 1.638 steps; CI95%: 0.384 to 2.892) and total time (MD: 2.274 seconds; CI95%: 0.531 to 4.017) to complete the Timed Up and Go test; and the step time (MD: 0.053; CI95%: 0.012 to 0.095; p=0.01) during walking.**

***Index Terms*— Inertial sensors, accidental falls, fall prediction, fall risk assessment, systematic literature review, meta-analysis**

I. Introduction

THE incidence of accidental falls among older adults, along with their impact in terms of morbidity and mortality, have turned them into a public health concern worldwide. It has been estimated that 28% to 45% of people aged 65 and over fall each year [1]. These events represent 18% to 40% of emergency department attendances and over 80% of all injury admissions to hospitals among the same age group. Among the most serious injuries resulting from falls are hip fracture and traumatic brain injury; the latter accounts for 46% of fatal falls among older adults [2].

Accidental falls have also a great impact in terms of costs for healthcare systems and for the society. Only in the United Kingdom, their annual cost to the National Health System has been estimated in £2.3 billion per year [3]. Moreover, falls lead to indirect costs, such as the loss of productivity of family members and other caregivers. The average lost earnings due to falls has been estimated in US$40,000 per year for the UK [1].

Nowadays, clinical fall risk assessment relies mostly on moderately to highly comprehensive medical, fall-risk specific and functional mobility assessment tools in the form of questionnaires, physical tests, gait analysis, and physical activity measurements [4]. Among the most popular assessment tests and tools are the Timed Up and Go (TUG) test [5], the Tinetti Assessment Tool [6], the STRATIFY score [7] and the Five-Times-Sit-to-Stand (FTSS) test [8].

More recently, researchers have investigated the potential use of instrumented fall-risk assessment and prediction tools based on features extracted from inertial sensors (i.e. accelerometers and gyroscopes) attached to the subject's body during specific assessment tasks (e.g. walking, quiet standing, sit-to-stand transitions) [9]–[11]. In those studies, machine learning methods were used to automatically identify fallers (F) and non-fallers (NF). Subjects were labelled as F/NF using at least one of the following methods: a fall-risk assessment test conducted in the clinical setting (e.g. TUG test), self-reported fall occurrence within a follow-up period from the assessment or fall history.

Howcroft *et al.* [9] and Shany *et al.* [11] have presented insightful accounts of features, classification models and validation strategies related to sensor-based fall-risk testing (SFRT). In their investigations, these authors found large heterogeneity in terms of sensor placement, tasks assessed, and sensor-based features. Not surprisingly, they also found disparate levels of reported sensitivity (55-100%), specificity (15-100%) and accuracy (62-100%).

To design effective interventions, it is crucial to identify the optimal combination of three factors: where to place the sensor, which task to be performed and which features should be extracted and analyzed. The latter is particularly relevant to

Manuscript received May 26, 2017; revised October 4, 2017; accepted October 30, 2017. L. Montesinos is sponsored by CONACyT (the Mexican National Council for Science and Technology). R. Castaldo is sponsored by the University of Warwick through an EPSRC IAA grant (EP/R511808/1).
(Corresponding author: Leandro Pecchia)

L. Montesinos (l.montesinos-silva@warwick.ac.uk), R. Castaldo (r.castaldo@warwick.ac.uk) and L. Pecchia (l.pecchia@warwick.ac.uk) are with the School of Engineering, University of Warwick, Coventry, CV4 7AL, UK. L. Montesinos is on leave from Tecnologico de Monterrey, Campus Ciudad de Mexico, Mexico.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2771383, IEEE Transactions on Neural Systems and Rehabilitation Engineering

overcome the limitations imposed by the curse of dimensionality (i.e. the difficulty and risk of training learning algorithms to discriminate between non-fallers and fallers in a high-dimensional feature space on the basis of a small pool of available data) [11]. Increasing the sample size would seem to be the logical solution. Unfortunately, achieving large sample sizes is one of the biggest challenges for this research area. Consequently, Shany *et al.* propose, as a more realistic solution, reducing the number of features prior to model building as sensibly as possible [11].

The goal of this systematic review and meta-analysis was to synthetize the empirical evidence regarding inertial sensor-based fall risk assessment and prediction in order to identify optimal combination of sensor placement, task and features aiming to support evidence-based design of new studies and real-life applications.

II. Methods

A. Literature Search

Potentially relevant articles on the risk assessment or prediction of falls based on features extracted from wearable inertial sensors were identified through a literature search in PubMed, EMBASE, IEEEXplore, Cochrane Central Register of Controlled Trials (CENTRAL), ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform electronic databases.

Articles were searched using Boolean combinations of the following keywords or equivalent Medical Subject Heading (MeSH) terms: accidental falls AND (risk assessment OR prediction) AND (sensor OR device OR wearable OR technology). No filter was applied at this stage.

Additional papers were identified performing a linear search along the references of relevant review articles previously published [9]–[12].

B. Inclusion and Exclusion Criteria

Papers were considered suitable for this review if they met all of the following criteria:

1) Original peer-reviewed journal articles published between January 2006 and December 2016 in English, Italian, Spanish or French languages (i.e. the languages on which the authors are qualified to understand a scientific text);
2) Studies in which the subjects were labelled as fallers and non-fallers (alternatively, high and low fall-risk), based on retrospective fall history, prospective fall occurrence, clinical assessment (e.g. the TUG test) or a combination of these methods;
3) A sample of at least 10 subjects with an average age of 60 or over;
4) Body-worn inertial sensors were used to characterize a physical task (e.g. walking or quiet standing) by extracting features from their signals, and;
5) Group statistics, specifically mean and standard deviation, for sensor-based features, as well as statistical significance level for the difference between groups were reported.

Papers were excluded if they included subjects with severe cognitive or motor impairment (e.g. Parkinson's disease, dementia).

Two authors independently assessed the suitability and methodological quality of the papers. A third author arbitrated when necessary.

C. Paper Selection and Data Extraction

Following the search strategy described above, all the records responding to the selected keywords were identified. After excluding duplicates (i.e. titles indexed in more than one database), studies were shortlisted according to inclusions/exclusions criteria by screening titles, abstracts and full-texts.

Subsequently, relevant data were extracted from the shortlisted studies; namely: first author and year of publication; number of participants and proportion of fallers; subject labelling method with details (e.g. follow-up period for prospective fall occurrence); type, quantity and placement of inertial sensors; test or task characterized via sensor-based features (e.g. the TUG test or quiet standing, respectively).

Finally, a listing of features reported in the shortlisted studies was compiled to enable further statistical analysis. For each feature the following items were included: name and category (i.e. linear acceleration, angular velocity, temporal, spatial, frequency, or non-linear features [9]), units, mean and standard deviation for each group (i.e. fallers and non-fallers), and trend over groups. A trend was represented with two arrows, ↓↓ (or ↑↑), if the mean value of a feature significantly ($p<0.05$) decreased (increased) for fallers compared to the mean value for non-fallers. Similarly, one arrow ↓ (or ↑) was used if the mean value of a feature non-significantly ($p>0.05$) decreased (increased) for fallers compared to the mean value for non-fallers. Sensor placement and assessed task for each feature were also included in the listing.

D. Statistical Analysis of Inertial Sensor-Based Features

Standard methods for the analysis of categorical data were applied on the feature listing with two objectives [13], [14]: 1) to investigate the level of association between trend significance status (i.e. non-significant or significant) and feature category, sensor placement and task, and; 2) to identify optimal triads of feature category, sensor placement and task.

Firstly, Pearson's chi-squared tests were performed in order to prove the association between trend significance status (dependent variable) and feature category, sensor placement and task (covariates). In other words, we aimed to prove that significant feature trends are dependent on feature category, sensor placement and/or task. A p-value < 0.05 was accepted as statistically significant evidence of a nonrandom association. Moreover, Pearson's Contingency (C) and Cramer's (V) coefficients were computed in order to quantify the level of association between each covariate and trend significance status. A C (V) coefficient of 0.1 (0.1), 0.287 (0.3) and 0.447 (0.5) were considered as evidence of small, medium and large level of association, respectively, as suggested in [15].

Secondly, significant triads of feature category, sensor placement and task were identified as follows. A three-way

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2771383, IEEE Transactions on Neural Systems and Rehabilitation Engineering

contingency table containing the abovementioned covariates was created using the subset of features containing only significant trends. Pearson residuals were computed for each triad in the table and used to characterize the strength (value) and nature (sign) of association for each triad. Large positive residuals are obtained when the observed frequency of significant features is substantially greater than the expected frequency, which would suggest significant features were more likely to arise from that specific triad. Conversely, large negative residuals are obtained when the observed frequency of significant features is substantially less than the expected, which would suggest significant features were less likely to arise from that specific triad. For interpretability, the following representation was used to report the results (instead of numerical values): two arrows, ↓↓ (or ↑↑), if the residuals were smaller (or larger) than -4 (or +4), revealing strong associations; one arrow, ↓ (or ↑), if the residuals were smaller (or larger) than -2 (or +2), revealing medium-associations, and; a dash, -, for residuals greater than or equal to -2 but smaller than or equal to +2, revealing weak associations. These thresholds are customarily used in the interpretation of Pearson residuals as a measure of strength of association [14]. A Pearson's chi-squared test of independence was performed to confirm the statistical significance of those associations (p-value < 0.05).

The software R version 3.2.3 was used to write the scripts to run this analysis.

E. *Meta-Analysis of Inertial Sensor-Based Features*

A meta-analysis of the features extracted from the shortlisted studies was conducted to identify significant individual features and their pivot values. Features were pooled for meta-analysis if: [feature was reported in at least two studies] AND [feature was computed for the same task/subtask] AND [sensor placement and type was the same across studies OR feature was independent of sensor placement and type (e.g. number of steps or stride time)]. Standard methods for combining and reporting continuous outcomes were employed to pool the features [16]: pooled sample size, mean difference (MD) with 95% confidence intervals (95% CI) and statistical significance level (p-value). MDs and 95% CIs were considered significant if the p-value was found to be smaller than 0.05.

Random or fixed effect models were selected based on heterogeneity across studies, assessed using the Q-statistic (computed via a Chi-squared test) and the I^2 statistic. A significant Q-statistic is indicative of dissimilar effect sizes across studies; a threshold significance level of 0.1 was selected as statistically significant value as suggested in [16]. The I^2 statistic indicates the percentage of the variability in effect sizes due to heterogeneity across studies, and not due to sampling error within studies. An I^2 from 30% to 60%, 50% to 90% and 75% to 100% represent moderate, substantial and considerable heterogeneity, respectively [17].

The R package meta_4.8-4 was used to conduct the meta-analysis [18]. The default options for both fixed and random models were used; i.e. the inverse variance method for study weighting and the DerSimonian-Laird estimate for the random effects model [19].

F. *Quality Appraisal of Shortlisted Studies*

The methodological quality of the studies was assessed using the checklist provided in the supplementary material (Document S1). This checklist was adapted from Downs and Black [20]. It contains 15 questions that are scored "yes" or "no/unclear". These questions are organized in 3 dimensions:

- Reporting (11 items) – which assessed whether the information provided in the paper was clear and sufficient to replicate the study and appraise its validity.
- External validity (2 items) –which addressed the extent to which the findings of the study could be generalized to a wider population and context.
- Internal validity (2 items) –which assessed whether the evidence at hand suggests that the study was designed and conducted to minimize bias and confounding.

A summary of the main findings is provided in this paper in an attempt to reveal the methodological issues that future studies in the field should address in order to produce more valid scientific evidence.

III. Results

According to the search strategy described above, 481 records were identified through database search and 18 through linear search. After removing 51 duplicates, 448 titles were screened by title and 257 were excluded as they did not meet the inclusion/exclusion. From the remaining 191 titles, 127 were removed after screening the abstract against

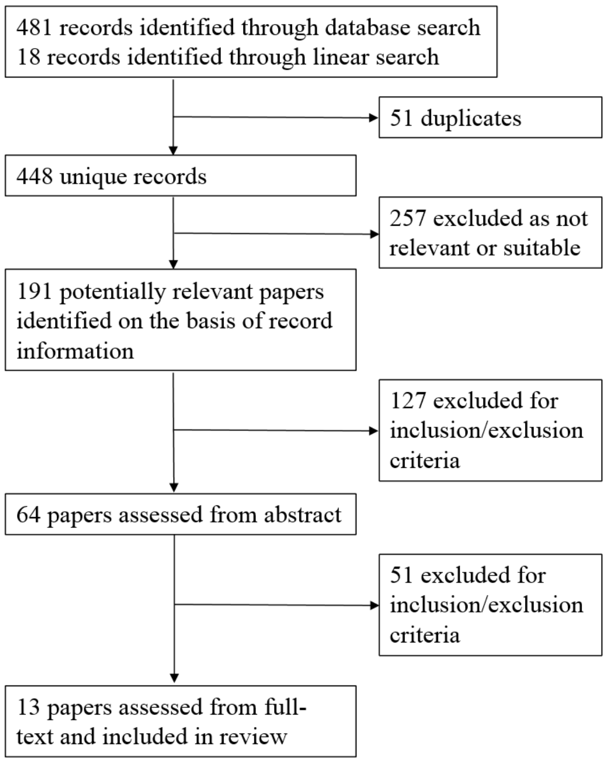

Fig. 1. Flowchart indicating the results of the systematic review with inclusions and exclusions.

TABLE I
DESCRIPTION OF SHORTLISTED STUDIES

Author, Year	Subjects (Fallers)	Mean age ± SD (years)	(Non-)Faller labelling method	Type of sensor	# of sensors	Location	Task
Kojima, 2008	153 (22)	71 ± 7.7	Retrospective fall history [b]	Accelerometer	1	Lower back	Walking
O'Sullivan, 2009	17 (12)	77 ± 7.5	Retrospective fall history [c]	Accelerometer	1	Lower back	Quiet standing
Greene, 2010	349 (207)	72.4 ± 7.4	Retrospective fall history [c]	Gyroscope	2	Shins	Timed Up and Go test
Paterson, 2011	97 (54)	68.7 ± 7.1	Prospective fall occurrence	Accelerometer	2	Feet	Walking
Weiss, 2011	41 (23)	78.2 ± 6.2	Retrospective fall history [b]	Accelerometer	1	Lower back	Timed Up and Go test
Doheny, 2012	40 (19)	71.4 ± 7.3	Retrospective fall history [c]	Accelerometer and Gyroscope	2	Shins / Lower back	Walking and Quiet standing
Greene, 2012	120 (65)	73.7 ± 5.8	Retrospective fall history [c]	Accelerometer and Gyroscope	1	Lower back	Quiet standing
Itoh, 2012	30 (7)	75 ± 5.7 [a]	Retrospective fall history	Accelerometer	4	Lower back, Knee, Ankle, Big toe	Walking
Senden, 2012	100 (50)	76.5 ± 5.7	Risk assessment tool	Accelerometer	1	Lower back	Walking
Doheny, 2013	39 (19)	71.5 ± 6.6	Retrospective fall history	Accelerometer	2	Thigh, Sternum	Five-Times Sit-to-Stand test
Doi, 2013	73 (16)	80.7 ± 7.8	Prospective fall occurrence	Accelerometer	2	Lower back, Upper back	10-m Walk test
Weiss, 2013	71 (32)	78.4 ± 4.7	Retrospective fall history [b]	Accelerometer	1	Lower back	Walking
Cui, 2014	81 (39)	78.4 ± 4.8	Retrospective fall history [b]	Accelerometer	1	Lower back	Walking

[a] Estimated from the data reported in paper
[b] Two or more falls within the recall period
[c] One or more falls within the recall period, or one fall resulting in injury or requiring medical attention

inclusion/exclusion criteria, which left 64 papers to be read in full-text. After reading the full-text, 51 were excluded due to inclusion/exclusion criteria. Therefore, 13 studies were shortlisted for this review [21]–[33]. A flowchart of the study selection process is shown in Fig. 1.

Importantly, there were some papers among the excluded ones which are noteworthy for the novelty of their approaches to the problem, their methodological quality and their results, but that failed to meet inclusion criterion 5. This is the case of the papers by Toebes et al [34] and Riva et al [35], who investigated the association between fall history and gait dynamic stability non-linear features (e.g. the Maximum Lyaponuv exponent, Multiscale entropy and Recurrence quantification analysis). Moreover, Rispens et al [36] and van Schooten et al [37] investigated the association of fall history and ambulatory (i.e. daily-life) gait measures. Finally, van Schooten et al [38] used survival models to describe the association of daily-life gait measures and prospective falls.

A. Characteristics of Shortlisted Studies

The 13 studies enrolled from 17 to 349 subjects each (mean ± standard deviation: 93.15 ± 86.18 subjects), for a cumulative population of 1,211. Overall, the studies included 565 fallers/high-risk subjects, i.e. 47% of the cumulative population. However, this proportion ranged from 14 to 71% across the 13 selected studies. The majority of studies (92%) included both men and women, except for one study which included only women [24]. Subjects were enrolled in a clinic as part of a larger clinical research project in 4 studies [23], [26], [27], [30], in a community center in 1 study [31], in a hospital's physiotherapy service in 1 study [22], and via letter sent to members of the community in 1 study [24]; details about the recruitment process were not provided in 6 studies [21], [25], [28], [29], [32], [33].

Additional details about the shortlisted studies are reported in Table I.

Subjects were labelled as (non-)fallers using retrospective fall history in 10 studies, with a recall period of one year for 8 studies and 5 years for 2 studies; prospective fall occurrence through a one-year follow-up period in 2 studies; and a clinical assessment tool (the Tinetti scale) in one study.

Tri-axial accelerometers and gyroscopes were the only type of inertial sensor used in 10 and 1 studies respectively; a combination of sensors were used in 2 studies. In 7 studies, only one sensor was used; in 5 studies two sensors were used; and 1 study used four sensors.

The most common sensor placement was the lower back (i.e. approximately on L3) with 10 studies, followed by shins and feet with 2 studies each. Other placements were knee, ankle, thigh, sternum and upper back (i.e. approximately on C7), with one study each. If placements are grouped in upper body (trunk) and lower body (lower limbs), there were eleven (91.7%) and seven (58.3%) studies, respectively.

Inertial signals were acquired during the following tasks: walking otherwise than a standardized test (7 studies), quiet standing (3 studies), the TUG test (2 studies), the 10-Meters Walking test (10MWT) (1 study), and the Five-Times Sit-to-Stand (FTSS) test (1 study). A brief description of these tasks is presented in Table II; for a more detailed description the reader may refer to the referenced paper.

B. Inertial Sensor-Based Features and Their Trends

The full listing of features extracted from inertial sensors that were reported in the 13 selected papers is provided as supplementary material (Table S1). Green *et al.* [23] reported features for all the subjects included in their analysis as well as for some subgroups separately (i.e. males, females < 75 year old and females ≥ 75 years old). However, only the results for all the subjects were included in this review. Moreover, Doheny *et al.* [26] performed an instrumented gait assessment four

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2771383, IEEE Transactions on Neural Systems and Rehabilitation Engineering

TABLE II
DESCRIPTION OF TASKS CHARACTERIZED USING INERTIAL SENSORS

Test / Task	Description
Walking	The individual is instructed to walk: ▪ 8 to 10 steps on a straight trajectory at comfortable and maximum speeds [21] ▪ 7 minutes at self-selected speed around a continuous walking circuit comprised by two straight sections (12 meters long) placed 3 meters apart [24] ▪ 3 meters at comfortable speed along a straight trajectory [26] ▪ 10 meters at a self-selected pace on a straight course which included stepping over six obstacles separated by 1.5 m [28] ▪ 20 meters on a straight course and back to the starting point at preferred speed [29] ▪ 1 minute or longer walking bouts during daily life activities [32] ▪ 1 minute under 3 different conditions: 1) baseline, usual walk; 2) baseline, usual walk with harness; 3) an obstacle course walk with harness [33]
Quiet standing	The individual was instructed to stand still for: ▪ 30 seconds with eyes open (EO), eyes closed (EC) and on a mat with eyes open (MAT EO) and closed (MAT EC) [22], [26] ▪ 40 seconds with eyes open (EO) in a semi tandem stance and 30 seconds with eyes closed (EC) [27]
Timed Up and Go test	The individual is instructed to rise from a chair, walk 3 meters at comfortable speed on a straight trajectory, turn around, walk back to the chair and sit down. [23], [25]
10-Meters Walking test	The individual is instructed to walk 10 meters at comfortable speed on a straight trajectory. A common practice is to use the intermediate 6 meters to allow for acceleration and deceleration. [31]
Five-Times-Sit-to-Stand test	The individual is instructed to keep her arms folded across her chest for the duration of the test and to fully stand up and sit back down five times as quick as possible. [30]

TABLE III
FREQUENCY TABLES FOR FEATURES BY TASK, SENSOR PLACEMENT AND FEATURE CATEGORY

	(A) All features (N=175)		(B) Significant features (N=84)	
	Count	%	Count	%
Task				
Walking [a]	110	62.9	61	72.6
Quiet standing	48	27.4	15	17.8
Sit-to-Stand / Stand-to-Sit [b]	14	8	5	6
TUG [c] test	3	1.7	3	3.6
Sensor placement				
Lower back	98	56	49	58.3
Shins	60	34.3	33	39.3
Foot	7	4	0	0
Sternum	4	2.3	0	0
Upper back	3	1.7	2	2.4
Knee	3	1.7	0	0
Feature category				
Linear acceleration	48	27.4	20	23.8
Temporal	45	25.7	19	22.6
Frequency	42	24	16	19
Angular velocity	32	18.3	25	29.8
Spatial	7	4	4	4.8
Non-linear	1	0.6	0	0

[a] Including walking tasks carried out as part of a test, e.g. the Timed Up and Go test
[b] Including sit-to-stand and stand-to-sit tasks carried out as part of a test, e.g. Timed Up and Go test
[c] Timed Up and Go test

times along the same day. However, only the results of the first assessment (between 9:00 and 9:30 am) were included in the review.

In summary, 93 distinct features were identified in the selected studies and categorized similarly to [9]: linear acceleration (15 features, 16.1%), angular velocity (28 features, 30.1%), spatial (4 features, 4.3%), temporal (24 features, 25.8%), frequency (21 features, 22.6%) and non-linear (1 feature, 1.1%).

These features were reported 175 times in the selected studies out of which 84 times (48%) they exhibited a significant trend.

TABLE IV
MEASURES OF ASSOCIATION BETWEEN FEATURE SIGNIFICANCE STATUS AND COVARIATES

Covariate	χ^2	p-value	C	V	Association level [a]
Task	11.94	**< 0.01**	0.253	0.261	Medium
Sensor placement	14.68	**0.01**	0.278	0.290	Medium
Feature category	15.82	**< 0.01**	0.288	0.301	Medium

χ^2: Pearson's chi-squared statistic for the association test in which the null hypothesis is 'no association'
C: Pearson's contingency coefficient
V: Cramer's coefficient
[a] A C (V) of 0.100 (0.1), 0.287 (0.3) and 0.447 (0.5) are considered as evidence of small, medium and large association, respectively

Table III summarizes the frequency of features per feature category, task and sensor placement for the complete listing of features (column A) and for the subset of features showing significant trends (column B).

C. *Statistical Analysis of Inertial Sensor-Based Features*

The results from the Pearson's chi-squared tests and the measures of association revealed statistically significant associations between feature significance and feature category, sensor placement and task (Table IV).

Furthermore, the computed Pearson residuals for the three-way table containing feature category, task and sensor placement as covariates revealed strong to very strong associations for 9 triads. Table V summarizes these results. As an example, the double arrow, '↑↑', for the triad 'angular velocity-walking-shins' means that significant features are much more likely to arise from this combination. Conversely, the single arrow, '↓', for the triad 'angular velocity-walking-lower back' means that significant features are less likely to arise from this combination. The '-' symbol indicated that the significance of a feature is not particularly affected by its category, sensor placement or task.

D. *Meta-Analysis of Inertial Sensor-Based Features*

Based on the selection criteria for the meta-analysis, 20 features were pooled using the methods described above. Table VI shows the trend and values for those features, as well as the number of subjects in each group. It also shows the task and the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2017.2771383, IEEE Transactions on Neural Systems and Rehabilitation Engineering

TABLE V
ASSOCIATION TREND AND STRENGTH FOR ALL POSSIBLE TRIADS OF FEATURE CATEGORY, TASK AND SENSOR PLACEMENT

Feature category	Quiet standing	SS	TUG	Walking	Sensor placement
Angular velocity	-	-	-	↓	Lower back
	-	-	-	↑↑	Shins
	-	-	-	-	Upper back
Frequency	-	-	-	↑	Lower back
	-	-	-	↓	Shins
	-	-	-	↑	Upper back
Linear acceleration	↑↑	↑↑	-	-	Lower back
	-	-	-	↓	Shins
	-	-	-	-	Upper back
Spatial	-	-	-	-	Lower back
	-	-	-	-	Shins
	-	-	-	-	Upper back
Temporal	-	-	-	-	Lower back
	-	-	↑	-	Shins
	-	-	-	-	Upper back

SS: Sit-to-Stand / Stand-to-Sit; TUG: Timed Up and Go test
↓↓ (↑↑): substantially stronger negative (positive) association for a specific triad of feature category, task and sensor placement
↓ (↑): strong negative (positive) association for a specific triad of feature category, task and sensor placement
-: either negative or positive non-significant association for a specific triad of feature category, task and sensor placement

sensor placement for each feature.

Linear acceleration features included in the meta-analysis were: Root Mean Square (RMS) value (expressed in g-force units) of acceleration signal in the mediolateral (ML) direction assessed at the lower back during quiet standing with both eyes open and eyes closed (ML RMS of acceleration). This feature is related to postural stability during standing.

Spatial features included in the meta-analysis were: number of steps during the Timed Up and Go (TUG) test, and step length as estimated from inertial signals measured during the walking stage of the TUG test or other walking task.

Temporal features included in the meta-analysis were: cadence (i.e. steps per minute); gait speed; step time; stance time; swing time; stride time; total time to complete the TUG test; single and double support time, i.e. the time during which only one foot and both feet are in contact with the walking surface, respectively, expressed as a percentage of a gait cycle; and the Coefficient of Variation (CV) for step, stance, swing, stride, single and double support times. The CV is the ratio of the standard deviation and mean for a given feature, expressed as a percentage; hence, it is a standardized measure of dispersion of the distribution of feature values.

All the spatial and temporal features included in the meta-analysis are widely used in clinical gait analysis [39].

One frequency feature was included in the meta-analysis: the Harmonic Ratio (HR) of trunk acceleration in the vertical (VT) direction. The HR has been defined as the ratio of even to odd signal harmonics extracted from the spectrum of the acceleration signal and has been suggested as a measure of the stability and smoothness of trunk movement during gait [31].

Neither angular velocity nor non-linear features were included in the meta-analysis, as none of them met the criteria to be pooled; i.e. either they were reported only in one study or they were measured during different tasks or at different sensor body placements.

The relative pooling weight of each study is reported in Table VI. The results of the pooling are reported in Table VII, where also the trend of the pooled features is shown.

Four out of twenty pooled features showed a statistically significant trend associated to fallers. A significantly higher RMS value for the ML acceleration signal (MD: 0.01 g; CI95%: 0.006 to 0.014; p<0.01) during quiet standing with eyes closed. Additionally, a significantly higher number of steps (MD: 1.638 steps; CI95%: 0.384 to 2.892; p=0.01) and a significantly higher total time to complete the TUG test (MD: 2.274 seconds; CI95%: 0.531 to 4.017; p<0.01). Finally, a significantly higher step time (MD: 0.053; CI95%: 0.012 to 0.095; p=0.01).

E. *Quality Appraisal of Shortlisted Studies*

All the studies reported aim of the study; experimental protocol (i.e. task, sensor quantity and placement); technical specifications of the sensor; methods for signal processing, feature extraction and statistical analysis; and features' summary statistics per group (non-fallers and fallers). However, only 7 studies reported actual p-values (e.g. 0.035 rather than <0.05) for the feature values' differences between groups [22], [25], [29]–[33].

Moreover, only 7 studies reported inclusion/exclusion criteria of participants and distribution of potential confounders per group (e.g. age and comorbidities) [24], [25], [27], [29], [31]–[33]. Therefore, the internal validity of 6 studies remains unclear, as unreported (or unobserved) variables could explain feature differences between fallers and non-fallers.

Finally, external validity was found for all shortlisted studies, as their samples were representative of the population under investigation and the task was representative of clinical fall-risk assessment protocols or daily-life activities.

IV. DISCUSSION

This systematic review analyzed the scientific literature focusing on the use of wearable inertial sensors for risk of fall assessment and prediction, exploring the sensitivity sensor-based features to sensor placement, task and feature category.

The statistical analysis of features reported in the 13 shortlisted studies revealed significant, very strong, positive associations in 3 different triads of feature category, task, and sensor placement:

- Angular velocity – Walking – Shins
- Linear acceleration – Quiet standing – Lower back
- Linear acceleration – Stand to sit/Sit to stand – Lower back

These results suggested that these are optimal combinations when using inertial sensors to discriminate between fallers and non-fallers. Other potentially good combinations, given their strong, positive associations are:

- Frequency – Walking - Lower back
- Frequency – Walking - Upper back
- Temporal - TUG - Shins

Conversely, our findings suggested that the use of following combinations should be avoided as they are less discriminative of fall status:

- Angular velocity – Walking - Lower back
- Frequency – Walking - Shins
- Linear acceleration – Walking - Shins

As for the meta-analysis, the results demonstrated that 4 features significantly increased ($p<0.05$) among fallers: the RMS acceleration in the mediolateral direction during quiet standing with eyes closed (MD: 0.01 g; CI95%: 0.006 to 0.014); the number of steps (MD: 1.638 steps; CI95%: 0.384 to 2.892) and total time (MD: 2.274 seconds; CI95%: 0.531 to 4.017) to complete the Timed Up and Go test; and the step time (MD: 0.053; CI95%: 0.012 to 0.095; p=0.01) during walking. These results suggest that these combinations of task and features may

TABLE VI
INERTIAL SENSOR-BASED FEATURES INCLUDED IN META-ANALYSIS

Feature (units)	Author, Year	Task	Sensor location	Trend	Weight (%)	Non-Fallers N	Non-Fallers Mean	Non-Fallers SD	Fallers N	Fallers Mean	Fallers SD
Linear acceleration features											
ML RMS acceleration (g)	Doheny, 2012	Quiet standing (EO)	Lower back	-	50.7	21	0.03	0.01	19	0.03	0.01
	Greene, 2012	Quiet standing (EO)	Lower back	↑↑	49.3	55	0.04	0.01	65	0.06	0.03
ML RMS acceleration (g)	Doheny, 2012	Quiet standing (EC)	Lower back	↑↑	44.3	21	0.03	0.01	19	0.04	0.01
	Greene, 2012	Quiet standing (EC)	Lower back	↑	55.7	55	0.04	0.01	65	0.05	0.02
Spatial features											
Number of steps (steps)	Weiss, 2011	TUG (Walking)	Lower back	↑	43.6	18	10.61	1.80	23	11.52	1.82
	Greene, 2010	TUG (Walking)	Shins	↑↑	56.4	142	10.60	2.40	207	12.80	3.80
Step length (m)	Weiss, 2011	TUG (Walking)	Lower back	↓	49.8	18	0.56	0.08	23	0.53	0.08
	Senden, 2012	Walking	Lower back	↑↑	50.2	50	0.51	0.13	50	0.66	0.09
Temporal features											
Cadence (steps/min)	Greene, 2010	TUG (Walking)	Shins	↓↓	50.2	142	108.00	19.30	207	99.20	19.30
	Senden, 2012	Walking	Lower back	↑↑	49.8	50	101.40	13.80	50	111.60	10.20
Gait speed (m/s)	Doi, 2013	10MWT	Lower back	↓↓	32.3	57	0.98	0.34	16	0.63	0.27
	Weiss, 2011	TUG (Walking)	Lower back	↓↓	34.1	18	0.68	0.10	23	0.60	0.09
	Senden, 2012	Walking	Lower back	↑↑	33.6	50	0.86	0.26	50	1.23	0.22
Step time (s)	Greene, 2010	TUG (Walking)	Shins	↑↑	26.5	142	0.60	0.10	207	0.70	0.10
	Weiss, 2011	TUG (Walking)	Lower back	↑↑	23.8	18	0.50	0.06	23	0.56	0.05
	Weiss, 2013	Walking	Lower back	↑↑	25	39	0.56	0.04	32	0.60	0.07
	Doheny, 2012	Walking	Shins	↑	24.6	21	0.57	0.05	19	0.58	0.05
Stance time (s)	Greene, 2010	TUG (Walking)	Shins	-	71.1	142	0.80	0.20	207	0.80	0.10
	Doheny, 2012	Walking	Shins	↑	28.9	21	0.68	0.10	19	0.70	0.08
Swing time (s)	Greene, 2010	TUG (Walking)	Shins	-	96.3	142	0.50	0.10	207	0.50	0.10
	Doheny, 2012	Walking	Shins	↓	3.7	21	0.47	0.25	19	0.43	0.04
Stride time (s)	Greene, 2010	TUG (Walking)	Shins	-	51.6	142	1.20	0.20	207	1.20	0.20
	Weiss, 2013	Walking	Lower back	↓↓	28.2	39	1.12	0.09	32	1.20	0.15
	Doheny, 2012	Walking	Shins	↑	20.2	21	1.11	0.11	19	1.13	0.11
Total time (s)	Weiss, 2011	TUG	Lower back	↑↑	52	18	8.68	1.62	23	10.10	1.61
	Greene, 2010	TUG	Shins	↑↑	48	142	12.40	5.10	207	15.60	6.50
Single support time (%)	Greene, 2010	TUG (Walking)	Shins	-	68.9	142	80.00	10.00	207	80.00	10.00
	Doheny, 2012	Walking	Shins	↓↓	31.1	21	78.39	5.59	19	75.53	4.67
Double support time (%)	Greene, 2010	TUG (Walking)	Shins	↓↓	55.4	142	50.00	20.00	207	40.00	20.00
	Doheny, 2012	Walking	Shins	↓	44.6	21	24.67	17.08	19	24.47	4.67
CV of step time (%)	Greene, 2010	TUG (Walking)	Shins	↑	43.3	142	40.30	22.90	207	42.00	21.00
	Doheny, 2012	Walking	Shins	↑	56.7	21	4.92	4.39	19	6.20	8.18
CV of stance time (%)	Greene, 2010	TUG (Walking)	Shins	↓	65.6	142	45.00	20.40	207	43.30	19.30
	Doheny, 2012	Walking	Shins	↑	34.4	21	6.03	8.67	19	7.40	10.16
CV of swing time (%)	Greene, 2010	TUG (Walking)	Shins	↓	43	142	31.00	22.00	207	28.10	19.90
	Doheny, 2012	Walking	Shins	↑↑	57	21	5.06	2.97	19	7.26	4.94
CV of stride time (%)	Greene, 2010	TUG (Walking)	Shins	↑	58.8	142	23.40	14.70	207	24.00	13.20
	Doheny, 2012	Walking	Shins	↑	41.2	21	4.19	5.56	19	4.96	6.01
CV of single support time (%)	Greene, 2010	TUG (Walking)	Shins	↑	38.8	142	21.10	19.20	207	22.90	15.70
	Doheny, 2012	Walking	Shins	↑	61.2	21	4.08	4.51	19	5.41	5.21
CV of double support time (%)	Greene, 2010	TUG (Walking)	Shins	↓	52.6	142	82.60	27.80	207	80.70	26.60
	Doheny, 2012	Walking	Shins	↑	47.4	21	10.02	9.61	19	16.54	12.39
Frequency features											
VT Harmonic ratio (n.u.)	Doi, 2013	10MWT	Lower back	↓↓	50.3	57	2.69	0.93	16	2.07	0.64
	Senden, 2012	Walking	Lower back	↑↑	49.7	50	2.18	1.09	50	3.09	1.25

EO: Eyes Open; EC: Eyes Closed; TUG: Timed Up and Go test; 10MWT: 10-Meters Walking Test
↓↓ (↑↑): significantly lower (higher) for subjects in the fallers (high-risk) subgroup
↓ (↑): lower (higher) for subjects in the fallers (high-risk) subgroup
-: No difference between subgroups

TABLE VII
POOLED INERTIAL SENSOR-BASED FEATURES

Feature (units)	Heterogeneity I^2 (%)	Q	p-value	Model	Subjects	Weighted Mean Difference MD	CI95%	p-value	Trend
Linear acceleration features									
ML RMS of acceleration, EO (g)	93.6	15.57	< 0.01	Random	160	0.010	(-0.001; 0.030)	0.32	↑
ML RMS of acceleration, EC (g)	0	0	1	Fixed	160	**0.010**	**(0.006; 0.014)**	**< 0.01**	↑↑
Spatial features									
Number of steps (steps)	73.9	3.83	0.05	Random	390	**1.638**	**(0.384; 2.892)**	**0.01**	↑↑
Step length (m)	96.5	28.58	< 0.01	Random	141	0.060	(-0.116; 0.237)	0.50	↑
Temporal features									
Cadence (steps/min)	97.1	35.01	< 0.01	Random	449	0.661	(-17.958; 19.281)	0.94	↑
Gait speed (m/s)	97.6	84.47	< 0.01	Random	214	-0.016	(-0.376; 0.345)	0.93	↓
Step time (s)	88.1	25.18	< 0.01	Random	501	**0.053**	**(0.012; 0.095)**	**0.01**	↑↑
Stance time (s)	0	0.35	0.55	Fixed	389	0.006	(-0.024; 0.036)	0.71	↑
Swing time (s)	0	0.5	0.48	Fixed	389	-0.002	(-0.022; 0.020)	0.89	↓
Stride time (s)	57.3	4.68	0.09	Fixed	460	0.026	(-0.004; 0.057)	0.09	↑
Total time (s)	79.6	4.91	0.03	Random	390	**2.274**	**(0.531; 4.017)**	**< 0.01**	↑↑
Single support time (%)	53.3	2.14	0.14	Fixed	389	-0.888	(-2.662; 0.885)	0.33	↓
Double support time (%)	79.4	4.85	0.03	Random	389	-5.625	(-15.174; 3.924)	0.25	↓
CV of step time (%)	0	0.02	0.90	Fixed	389	1.462	(-1.649; 4.572)	0.36	↑
CV of stance time (%)	0	0.69	0.41	Fixed	389	-0.643	(-4.095; 2.809)	0.71	↓
CV of swing time (%)	73	3.7	0.05	Random	389	0.005	(-4.945; 4.954)	1	↑
CV of stride time (%)	0	0.01	0.94	Fixed	389	0.670	(-1.641; 2.981)	0.57	↑
CV of single support time (%)	0	0.04	0.85	Fixed	389	1.512	(-0.863; 3.887)	0.21	↑
CV of double support time (%)	69.9	3.32	0.07	Random	389	2.095	(-6.146; 10.336)	0.62	↑
Frequency features									
VT Harmonic ratio (n.u.)	95.9	24.44	< 0.01	Random	173	0.140	(-1.359; 1.640)	0.85	↑

EO: Eyes Open; EC: Eyes Closed
↓↓ (↑↑): significantly lower (higher) for subjects in the fallers (high-risk) subgroup
↓ (↑): lower (higher) for subjects in the fallers (high-risk) subgroup
MD: Mean difference; CI95%: Confidence Interval at 95%; n.u.: dimensionless
Bold values indicate statistically significant trends ($p < 0.05$)

be useful more effective for fall risk assessment.

Additionally, 5 features exhibited a consistent trend across the selected studies. These features were: step time, CV for step time, CV for stride time and CV for single support time, which showed a higher value for fallers when compared to non-fallers; and double support time, which showed a lower value for the same group. However, these trends were not found statistically significant when pooled in the meta-analysis. It may be explained by the high values of standard deviation reported by Green et al [23], which was included in the pooling for these features. No clear explanation for such variability within that study can be inferred from the paper.

In contrast, 7 features showed an opposite trend across the selected studies: step length, cadence, gait speed, harmonic ratio in the vertical direction, CV for stance time, CV for swing time, and CV double support time. Importantly, for 4 of these features the methods used to classify subjects as (non-)fallers were also inconsistent between studies: step length and cadence were pooled from [25] and [29], in which the classification methods were retrospective fall history and fall risk assessment tool (Tinetti scale), respectively. Gait speed was pooled from [25], [28] and [31], the latter adding prospective fall occurrence to the diversity of classification methods. Finally, harmonic ratio on the vertical direction was pooled from [29] and [31], combining subjects classified as fallers via two different methods as well. This fact may represent an important source of between-study heterogeneity, as reflected by the high values of I^2 (>95%) and low significance levels ($p < 0.01$) obtained in the heterogeneity test for these features. Unfortunately, the low number of studies reporting on the same feature made unfeasible to explore possible sources of heterogeneity using quantitative approaches (e.g. via subgroup analysis stratified by study and/or patient characteristics).

Moreover, 5 features showed an ambiguous trend across the selected studies, as they were reported with no mean difference between non-fallers and fallers in one study, while exhibiting a trend (significant or not) in another study. These features were: the RMS value of acceleration in the mediolateral direction with eyes open, and stance time, swing time, stride time, and single support time during walking.

All in all, the evidence gathered in this review suggests that assessing the Timed Up and Go test using wearable sensors located on the shins through angular velocity, temporal (e.g. total time and step time) and spatial (e.g. number of steps) features may represent an optimal combination to discriminate fallers from non-fallers. Additionally, the triad "linear acceleration-quiet standing-lower back" seems to be a sensible choice as well.

Nevertheless, it should be stressed that these results are limited, as they are based only on features reported in the 13 papers included in the review. Hence, they are unable to provide a representative inference of all features used and all studies published, but not included in the review. It means that there might be some other sensor-based features that are discriminant between non-fallers and fallers but were not included in this systematic review as they were not reported as required by the inclusion criteria. This may be the case of some features reported in [34]–[38].

Finally, a comment regarding heterogeneity in "hit rate" (i.e. the ratio of all features to significant features expressed as a percentage) reported in the shortlisted studies is deemed relevant to this review. In some studies reporting a relatively high number of features (i.e. 28 or more) a hit rate ranging from 25 to 66% was achieved [23], [26], [27]. In contrast, some studies reporting a low number of features (i.e. 7 or less) achieved hit rates above 85%, with two studies reporting a surprising 100% [29], [31], [33]. From these studies, it was not clear if the authors investigated a low number of features or if they investigated a large number of features but only reported the most significant ones. Even if reporting bias (a.k.a. selective reporting) should not be concluded from this finding, it should at least make us aware of the potential presence of this practice in our field. This practice could undermine the findings of future studies, making more difficult to converge to meaningful conclusions.

V. Conclusions

In conclusion, this paper demonstrated that there are high and significant interactions among sensor placement, task and feature category to assess the risk of falling. This systematic review provided a framework for future study design, highlighting dependences among those factors. In addition, the review generated a comprehensive inventory of the features reported so far from inertial sensors for fall risk assessment in older adults, summarizing their trends and whether these were found statistically significant or not in each study. The statistical analysis of those features demonstrated that the triad 'angular velocity-walking-shins' has shown more discriminative power between non-fallers and fallers than others. Finally, the meta-analysis demonstrated that 4 features resulted significantly different between non-fallers and fallers. However, most features were not included in the meta-analysis because they were not reported with sufficient homogeneity in at least 2 studies, suggesting that future studies are required to produce more evidence that allows to conduct a more comprehensive meta-analysis. Future studies should consider the evidence resulting from our review, in particular for: 1) the selection of the features to be further explored; 2) the sensor placement; and 3) the task used to assess the risk of falling. Those studies could also benefit from the adoption of some practices more common in clinical research, such as the definition of participant inclusion/exclusion criteria, inclusion of potential confounders in the analysis and ultimately the ex-ante publication of the full study protocol prior to the study conduction. These practices aim to reduce the risk of bias and confounding, thus giving more validity to the study.

Acknowledgments

This study brought together existing data extracted from 13 original papers included in the review and cited in the reference section. These data are provided in full as supplementary material accompanying this paper.

The authors would like to thank Miss Pasqua Meccariello for her valuable contributions to the initial steps of this work, carried out during her time as visiting researcher at the School of Engineering, University of Warwick.

References

[1] World Health Organization, "WHO global report on falls prevention in older age." World Health Organization, 2008.

[2] N. M. Peel, "Epidemiology of Falls in Older Age," *Can. J. Aging Rev. Can. Vieil.*, vol. 30, no. 01, pp. 7–19, Mar. 2011.

[3] "Falls in older people: assessing risk and prevention | Guidance and guidelines | NICE." [Online]. Available: https://www.nice.org.uk/guidance/cg161?unlid=297510847201713154334. [Accessed: 26-Jan-2017].

[4] K. L. Perell, A. Nelson, R. L. Goldman, S. L. Luther, N. Prieto-Lewis, and L. Z. Rubenstein, "Fall risk assessment measures an analytic review," *J. Gerontol. A. Biol. Sci. Med. Sci.*, vol. 56, no. 12, pp. M761–M766, 2001.

[5] S. Mathias, U. S. Nayak, and B. Isaacs, "Balance in elderly patients: the 'get-up and go' test," *Arch. Phys. Med. Rehabil.*, vol. 67, no. 6, pp. 387–389, Jun. 1986.

[6] M. E. Tinetti, T. F. Williams, and R. Mayewski, "Fall risk index for elderly patients based on number of chronic disabilities," *Am. J. Med.*, vol. 80, no. 3, pp. 429–434, Mar. 1986.

[7] D. Oliver, M. Britton, P. Seed, F. C. Martin, and A. H. Hopper, "Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies," *BMJ*, vol. 315, no. 7115, pp. 1049–1053, Oct. 1997.

[8] M. Csuka and D. J. McCarty, "Simple method for measurement of lower extremity muscle strength," *Am. J. Med.*, vol. 78, no. 1, pp. 77–81, Jan. 1985.

[9] J. Howcroft, J. Kofman, and E. D. Lemaire, "Review of fall risk assessment in geriatric populations using inertial sensors," *J. NeuroEngineering Rehabil.*, vol. 10, no. 1, p. 91, 2013.

[10] A. Ejupi, S. R. Lord, and K. Delbaere, "New methods for fall risk prediction," *Curr. Opin. Clin. Nutr. Metab. Care*, vol. 17, no. 5, pp. 407–411, Sep. 2014.

[11] T. Shany, K. Wang, Y. Liu, N. H. Lovell, and S. J. Redmond, "Review: Are we stumbling in our quest to find the best predictor? Over-optimism in sensor-based models for predicting falls in older adults," *Healthc. Technol. Lett.*, vol. 2, no. 4, pp. 79–88, Aug. 2015.

[12] D. Hamacher, N. B. Singh, J. H. Van Dieen, M. O. Heller, and W. R. Taylor, "Kinematic measures for assessing gait stability in elderly individuals: a systematic review," *J. R. Soc. Interface*, vol. 8, no. 65, pp. 1682–1698, Dec. 2011.

[13] B. Vidakovic, *Statistics for bioengineering sciences: with MATLAB and WinBUGS support*. New York: Springer, 2011.

[14] M. Friendly and D. Meyer, *Discrete data analysis with R: visualization and modeling techniques for categorical and count data*. Boca Raton: CRC Press, Taylor & Francis Group, 2016.

[15] J. Cohen, *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ ; London: L. Erlbaum Associates, 1988.

[16] A. J. Sutton, Ed., *Methods for meta-analysis in medical research*. Chichester: Wiley, 2000.

[17] "Cochrane Handbook for Systematic Reviews of Interventions | Cochrane Training." [Online]. Available: http://training.cochrane.org/handbook. [Accessed: 26-Jan-2017].

[18] G. Schwarzer, J. R. Carpenter, and G. Rücker, *Meta-Analysis with R*. Cham: Springer International Publishing, 2015.

[19] R. DerSimonian and N. Laird, "Meta-analysis in clinical trials," *Control. Clin. Trials*, vol. 7, no. 3, pp. 177–188, Sep. 1986.

[20] S. H. Downs and N. Black, "The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions.," *J. Epidemiol. Community Health*, vol. 52, no. 6, pp. 377–384, 1998.

[21] M. Kojima, S. Obuchi, O. Henmi, and N. Ikeda, "Comparison of Smoothness during Gait between Community Dwelling Elderly Fallers and Non-Fallers Using Power Spectrum Entropy of

Acceleration Time-Series," *J. Phys. Ther. Sci.*, vol. 20, no. 4, pp. 243–248, 2008.

[22] M. O'Sullivan, C. Blake, C. Cunningham, G. Boyle, and C. Finucane, "Correlation of accelerometry with clinical balance tests in older fallers and non-fallers," *Age Ageing*, vol. 38, no. 3, pp. 308–313, Feb. 2009.

[23] B. R. Greene, A. O. Donovan, R. Romero-Ortuno, L. Cogan, C. Ni Scanaill, and R. A. Kenny, "Quantitative Falls Risk Assessment Using the Timed Up and Go Test," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 12, pp. 2918–2926, Dec. 2010.

[24] K. Paterson, K. Hill, and N. Lythgo, "Stride dynamics, gait variability and prospective falls risk in active community dwelling older women," *Gait Posture*, vol. 33, no. 2, pp. 251–255, Feb. 2011.

[25] A. Weiss, T. Herman, M. Plotnik, M. Brozgol, N. Giladi, and J. M. Hausdorff, "An instrumented timed up and go: the added value of an accelerometer for identifying fall risk in idiopathic fallers," *Physiol. Meas.*, vol. 32, no. 12, pp. 2003–2018, Dec. 2011.

[26] E. P. Doheny, B. R. Greene, T. Foran, C. Cunningham, C. W. Fan, and R. A. Kenny, "Diurnal variations in the outcomes of instrumented gait and quiet standing balance assessments and their association with falls history," *Physiol. Meas.*, vol. 33, no. 3, pp. 361–373, Mar. 2012.

[27] B. R. Greene *et al.*, "Quantitative falls risk estimation through multi-sensor assessment of standing balance," *Physiol. Meas.*, vol. 33, no. 12, pp. 2049–2063, Dec. 2012.

[28] T. Itoh *et al.*, "Development of a new instrument for evaluating leg motions using acceleration sensors (II)," *Environ. Health Prev. Med.*, vol. 17, no. 3, pp. 205–212, May 2012.

[29] R. Senden, H. H. C. M. Savelberg, B. Grimm, I. C. Heyligers, and K. Meijer, "Accelerometry-based gait analysis, an additional objective approach to screen subjects at risk for falling," *Gait Posture*, vol. 36, no. 2, pp. 296–300, Jun. 2012.

[30] E. P. Doheny *et al.*, "Falls classification using tri-axial accelerometers during the five-times-sit-to-stand test," *Gait Posture*, vol. 38, no. 4, pp. 1021–1025, Sep. 2013.

[31] T. Doi, S. Hirata, R. Ono, K. Tsutsumimoto, S. Misu, and H. Ando, "The harmonic ratio of trunk acceleration predicts falling among older people: results of a 1-year prospective study," *J. NeuroEngineering Rehabil.*, vol. 10, no. 1, p. 7, 2013.

[32] A. Weiss *et al.*, "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*, vol. 27, no. 8, pp. 742–752, 2013.

[33] X. Cui, C.-K. Peng, M. D. Costa, A. Weiss, A. L. Goldberger, and J. M. Hausdorff, "Development of a new approach to quantifying stepping stability using ensemble empirical mode decomposition," *Gait Posture*, vol. 39, no. 1, pp. 495–500, Jan. 2014.

[34] M. J. P. Toebes, M. J. M. Hoozemans, R. Furrer, J. Dekker, and J. H. van Dieën, "Local dynamic stability and variability of gait are associated with fall history in elderly subjects," *Gait Posture*, vol. 36, no. 3, pp. 527–531, Jul. 2012.

[35] F. Riva, M. J. P. Toebes, M. Pijnappels, R. Stagni, and J. H. van Dieën, "Estimating fall risk with inertial sensors using gait stability measures that do not require step detection," *Gait Posture*, vol. 38, no. 2, pp. 170–174, Jun. 2013.

[36] S. M. Rispens, K. S. van Schooten, M. Pijnappels, A. Daffertshofer, P. J. Beek, and J. H. van Dieën, "Identification of Fall Risk Predictors in Daily Life Measurements: Gait Characteristics' Reliability and Association With Self-reported Fall History," *Neurorehabil. Neural Repair*, vol. 29, no. 1, pp. 54–61, Jan. 2015.

[37] K. S. van Schooten, M. Pijnappels, S. M. Rispens, P. J. M. Elders, P. Lips, and J. H. van Dieën, "Ambulatory Fall-Risk Assessment: Amount and Quality of Daily-Life Gait Predict Falls in Older Adults," *J. Gerontol. A. Biol. Sci. Med. Sci.*, vol. 70, no. 5, pp. 608–615, May 2015.

[38] K. S. van Schooten *et al.*, "Daily-life gait quality as predictor of falls in older people: a 1-year prospective cohort study," *PLoS One*, vol. 11, no. 7, p. e0158623, 2016.

[39] R. Baker and H. M. Hart, *Measuring walking: a handbook of clinical gait analysis*. London: Mac Keith Press, 2013.