A comparison of cross-sectional and prospective algorithms for falls risk assessment

Barry R. Greene, *Senior IEEE Member*, Denise McGrath, Brian Caulfield, *IEEE Member*

Abstract— **Falls are the most common cause of injury and hospitalization and one of the principal causes of death and disability in older adults worldwide. Accurate identification of patients at risk of falls could lead to timely medical intervention, reducing the incidence of falls related injuries along with associated costs. The current best practice for studies of falls and falls risk recommends the use of prospective follow-up data. However, the majority of studies reporting sensor based methods for assessment of falls risk employ cross-sectional falls data (falls history). The purpose of this study was to compare the performance of sensor based falls risk assessment algorithms derived from cross-sectional (N=909) and prospective (N=259) datasets in terms of false positive rate. The utility of any classification algorithm is clearly limited by a high false positive rate. An estimate of the false positive rate for both cross-sectional and prospective algorithms was determined using an inertial sensor data set of 611 TUG tests from 55 healthy control subjects, with no history of falls. We aimed to determine which falls risk assessment algorithm is more effective at classifying falls risk in healthy control subjects. The cross-sectional algorithm correctly classified 94.11% of tests, while the prospective algorithm, correctly classified 79.38% of tests. Results suggest that sensor based falls risk assessment algorithms generated using cross-sectional falls data, may be more effective than those generated using prospective data in classifying healthy controls and reducing associated false positives.**

I. Introduction

The prevention of falls in older adults has become an increasingly important clinical challenge as the world's population continues to age. Approximately one third of adults over 65 years of age fall each year [1]. In the United States in the year 2000 alone, fatal falls cost $179 million, while non-fatal falls cost $19 billion. As the worldwide population ages, the incidence of falls and their associated costs are set to increase [2, 3]. Accurate identification of those patients at high risk of falls would facilitate appropriate and timely intervention, and could lead to improved quality of care and reduced associated hospital costs, due to reduced admissions and reduced severity of falls.

This study was completed as part of a wider programme of research within the TRIL Centre, (Technology Research for Independent Living). The TRIL Centre is a multi-disciplinary research centre, bringing together researchers from UCD, TCD and Intel, funded by Intel, GE and IDA Ireland. http://www.trilcentre.org.

Author BRG is supported by an Enterprise Ireland commercialization grant (CF/2013/3600).

B.R Greene is with the TRIL centre, University college Dublin, Ireland and Kinesis Health Technologies, Dublin, Ireland (e-mail: barry.greene@ucd.ie).

D. McGrath, School of Public Health, Physiotherapy and Population Science, University College Dublin, Ireland (email: denise.mcgrath@ucd.ie).

B. Caulfield, Insight Centre and School of Physiotherapy and Performance Science, University College Dublin, Ireland (e-mail: b.caulfield@ucd.ie).

Falls risk is generally assessed in a clinical setting by physiotherapists, geriatricians, clinical nurse specialists or occupational therapists. A variety of validated clinical recommendations exist for assessing falls risk [4-6]. However, these can be subjective, variable in administration and may require specialist expertise. An objective method for assessing falls risk, suitable both for use by non-experts and for deployment in a community care setting may find clinical application for screening and targeting of individuals at high risk of falls.

The current best practice for the use of outcome data in falls injury prevention trials, as recommended by the consensus guidelines [7], is the use of prospective fall diaries. This is because retrospective falls history data are considered to suffer from a number of limitations, namely: participant recall bias and gait changes due to previous falls. It is not currently clear which type of outcome measures would in reality, form the more reliable index of a patients' current falls risk, i.e., a person's falls risk at the time of assessment.

Recent research has reported a wide variety of inertial sensor based methods for assessing falls risk in older adults [8-13]. The majority of studies have validated their methods using cross-sectional (retrospective) falls data, i.e. each participant's history of falls prior to assessment [14, 15]. A smaller number of trials have reported a prospective validation of their methods, i.e. have followed-up participants for falls, for a period of time after assessment and used these falls data to validate their methods [10, 16].

To date no direct comparison has been made between cross-sectional and prospective methods for assessment of falls, nor to our knowledge has any validation been performed using healthy control data. The present authors have published both cross-sectional (employing falls history)[8] and prospective (employing falls follow-up data)[16] studies on falls risk assessment. We sought to determine which method would be most effective at classifying an independent set of healthy control data. For any real world deployment of a falls risk assessment algorithm, quantification of the false positive rate is very important, given that any system prone to high levels of false positives would quickly erode the confidence of the clinician interpreting the results. In the absence of an additional independent prospective falls test set, testing a falls risk assessment algorithm on an independent set of healthy control data is a method for determining an estimate of the falls positive rate under real world conditions.

II. METHOD

The 'timed up and go' (TUG) test is a standard mobility assessment used to screen for balance problems in older people [17-19]. The TUG test consists of the participant getting up from a chair, walking three metres, turning at a designated spot, returning to the seat and sitting down. The time taken to perform the test is recorded using a stopwatch. Current clinical practice suggests that elders with longer TUG times are more likely to fall than those with shorter times. The performances of older adults prone to falling can be very different from those who do not fall. Consequently, the TUG test is one of the most widely used tools for identifying elders at risk of falls [19] and has been recommended by the American Geriatrics Society/British Geriatrics Society (AGS/BGS) guidelines as a screening tool for identifying older people at increased risk of falls [20].

Nine hundred and nine TUG tests were obtained from participants assessed as part of the TRIL project (www.trilcentre.org), a large ageing research project. All participants provided informed consent and had their mobility assessed using the TUG test instrumented with inertial sensors in the TRIL clinic, St James's hospital (SJH), Dublin, Ireland between 2007 and 2012. Participants were at least 60 years old, had never experienced a stroke, and were able to walk without assistance. Ethical approval was received from the local research ethics committee.

Each participant's history of falls in the past 5 years was obtained by means of a questionnaire. A fall was defined as an event which resulted in a person coming to rest on the lower level regardless of whether an injury was sustained, and not as a result of a major intrinsic event or overwhelming hazard [21]. Falls outcome data were verified using collateral history from relatives as well as comparison with hospital records. All participants had a detailed clinical assessment and falls history consistent with the AGS guidelines [22]. Each participant completed at least one TUG test. During each TUG test, inertial sensor data were obtained from shank-mounted tri-axial gyroscopes in order to quantify each test (see Fig. 1).

A. *Healthy control data*

Two data sets of normal healthy controls were also collected under laboratory conditions, consisting of 503 TUG tests from 32 participants (control data set one (CS1)) and 108 TUG tests from 23 participants (control data set two (CS2)). All controls had no history of falls. Participants in CD1 (14 M, 18 F) had a mean age of 59.8±2.7yrs and a mean height and weight of 167.2±7.5cm and 74.3±13.6kg. Participants in CD2 (15 M, 8 F) completed 4-6 TUG trials each and had a mean age of 43.7±15.5, while their mean height and weight were 171.0±19.5cm and 79.7±25.7kg respectively. Sample sensor data for a healthy control subject is shown in Fig.1.

B. *Cross-sectional data*

Data derived from the 909 TUG tests obtained from four different waves assessed in the TRIL clinic were included in the analysis. This combined dataset consisted of one TUG trial per participant (292 M, 617 F). The falls history data for each participant was used as the outcome data for the cross-sectional falls risk assessment model, participants with a history of falls were termed 'fallers' while those with no history of falls were termed 'non-fallers'. The mean age of this cohort was 73.6±6.9, while the mean weight and height were 73.9±14.3kg and 165.2±9.2cm respectively.

C. *Prospective data*

Participants from a single wave assessed in the TRIL clinic were followed up for falls for a period of two years. Subsequent to baseline assessment, 259 (76 M, 183 F) participants were contacted by telephone approximately two years subsequent to their initial assessment and asked to complete a survey on their falls history subsequent to their initial assessment. The falls follow-up data obtained was used as the falls outcome data for the prospective falls risk assessment model. Participants that fell during the follow-up period were termed 'fallers', while those that did not fall were termed 'non-fallers'. The mean age of the cohort at the time of initial evaluation was 71.5±6.7 years, while the mean height and weight were 165.4±9.4 cm and 73.6±14.3 kg respectively.

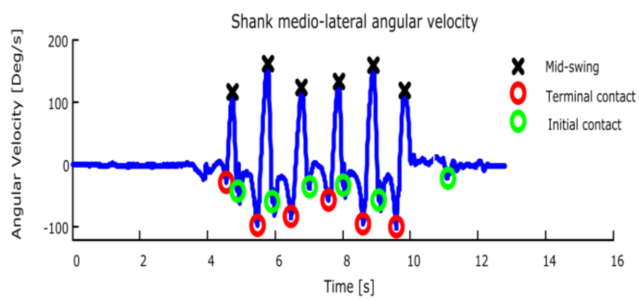

Figure 1: Medio-lateral shank angular velocity signal obtained from healthy control subject while performing a TUG test. Initial and terminal contact points (IC and TC) as well as mid-swing points are indicated.

D. *Data acquisition*

Each TUG test was conducted as follows; participants stood from a 46cm high chair with armrests, walked 3 metres, turned 180 degrees, and returned to the chair and sat down (see Fig. 2). They were instructed to perform this as fast as was safely possible. The clinician said 'go' when they started the recording, and ended the recording when the participant was re-seated with their back touching the back of the chair.

Inertial sensors (Shimmer research, Dublin, Ireland), containing tri-axial accelerometers and a tri-axial gyroscope, were used to quantify movement during each assessment. Participants wore inertial sensors on the anterior aspect of each shin (IS1 and IS2), with one axis aligned with the tibial bone, secured using elasticated tubular bandages (Tubi-grip). Inertial sensor data were sampled at 102.4Hz. Inertial sensor data were synchronously acquired in real-time via Bluetooth. Data were automatically saved to text format for subsequent offline analysis. Post-processing and analysis were conducted off-line using Matlab (Mathworks, VA, USA). Gyroscope data were calibrated using a published procedure [23] and low-pass filtered using a Butterworth IIR filter with a 20Hz corner frequency. 52 features quantifying the temporal, spatial, turning and rotational characteristics (details appear elsewhere [8, 16]) were extracted from inertial sensor data for each TUG test.

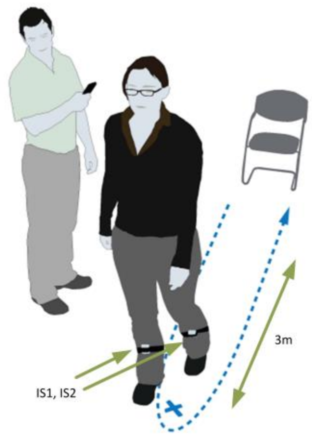

Figure 2. Experimental setup for quantifying TUG test using shank mounted inertial sensors (IS 1 and IS2).

D. Statistical analysis

Both the cross-sectional and prospective data sets were classified using a regularized discriminant classifier model, with regularization parameter values set to λ=0.1 and r=0.1 prior to analysis. The generalized performance of each model was evaluated using 10 repetitions of 10-fold cross validation [24], where training and testing sets are randomly selected for each repetition to obtain an unbiased estimate of generalized classifier performance. Features were selected using sequential forward feature selection [25] within the cross-validation procedure.

The performance of the classifier model was evaluated using a number of standard performance measures; the classification accuracy (Acc), defined as the percentage of participants correctly classified by the system as being a faller or non-faller. The sensitivity (Sens) is defined as the percentage of the faller class classified correctly. Similarly, specificity (Spec) is defined as the percentage of the non-faller class correctly identified as such by the system. Positive and negative predictive values were also calculated to provide a measure of the predictive power of positive and negative classifications. The positive predictive value (PPV) is defined as the proportion of participants, classified as fallers by the algorithm, who are correctly classified. Similarly, the negative predictive value (NPV) is the proportion of participants, classified as non-fallers by the algorithm, who are correctly classified. The area under the receiver operator characteristic curve (ROC), calculated using the discriminant function output is reported as an index of class discrimination. The values reported for each classifier performance metric were averaged across all cross-validation folds and repetitions. In order to ensure that the cross-sectional and prospective models were generated using the same method, the method for model generation reported here differ slightly from those reported in the original studies [8, 16].

III. Results

1) Comparing cross-sectional and prospective model

The performance of each algorithm in assessing falls in their respective cohorts is detailed in Table 1. For the cross sectional model, cross-validation yielded a mean classification accuracy of 70.02% in classifying participants according to falls history. Similarly, the prospective model was 76.19% accurate in classifying participants according to falls in the two years subsequent to assessment.

TABLE I. Cross validated estimates of falls risk assessment performance for cross-sectional and prospective falls risk models. Cross- sectional model uses falls history as outcome data while the prospective model uses falls follow-up data.

	Cross-sectional (N=909)	**Prospective (N=259)**
Acc (%)	70.02	76.27
Sens (%)	47.73	57.20
Spec (%)	84.72	83.63
PPV (%)	70.14	59.86
NPV (%)	69.19	82.54
ROC	0.67	0.69

2) Validating models using healthy control data

The CS1 and CS2 data sets were used to validate the performance of both classifier models in classifying healthy control subjects with no history of falls. The 'final' cross-sectional and prospective models were obtained by training both classifier models using the features selected through the cross-validated feature selection procedure, and all available data for each model. The performance of each classifier model in classifying the TUG test sensor data in CS1 and CS2 is detailed in Table 2.

TABLE II. Performance of cross-sectional and prospective falls risk models on two data sets of healthy control subjects (CD1 and CD2). CS1 contains 503 TUG tests from 32 participants while CS2 contains 108 TUG tests from 23 participants.

	Cross sectional	**Prospective**	**(N)**
Total correct (CS1)	497	404	503
Total correct (CS2)	79	81	108
Total correct	575	485	611
Acc (%)	94.11	79.38	-

IV. Discussion

Our results suggest that a cross-sectional falls risk assessment algorithm outperforms a prospective falls risk assessment algorithm in classifying healthy control subjects (94.11% compared to 79.38%). This means that the false positive rate for the cross-sectional algorithm was 5.89% compared to 20.62% for the prospective algorithm. This is despite higher reported values for classification accuracy for the prospective model compared to the cross-sectional model, on their respective cohorts (76.27% versus 70.02%). The cross-sectional data set was far larger (909 compared to 259) and contained a wider variety of gait patterns. This heterogeneity may have led to lower performance estimates for generalized performance compared to the prospective model and a previously published cross-sectional method [16], which was based on data from a single wave.

We made a key assumption in this study: we assumed that the healthy controls classified as fallers were *falsely* classified as such. While all control subjects were healthy and had no history of falls, prospective follow-up data was not available. However, given their relatively young age (mean 59.8±2.7 and 43.7±15.5 for CD1 and CD2 respectively) and robust physical state, is it reasonable to assume that all participants would register a low falls risk score, from a falls risk assessment algorithm that was robust to unseen data.

This study raises the following questions: 1) in falls risk research, where we are seeking the most accurate ways to identify those individuals who are at risk of falls, are we aiming to provide healthcare providers with the best tools to identify falls risk in the short-term, the medium-term or long-term?; 2) what type of data should we be using to validate our algorithms i.e. cross-sectional, prospective or both, for each of these different timeframes? In our study, the follow-up period for the prospective data set was two years. It is conceivable that the gait characteristics associated with falls may not have been present in all participants at the initial assessment and participants may have only developed these characteristics in the subsequent two years. In such a situation, it is arguable that falls history may represent a better index of a subjects' *current* falls risk.

Acknowledgment

We acknowledge the help and support of the staff of the TRIL Clinic, St James, hospital, Dublin and the participants involved in this study.

References

[1] T. Masud and R. O. Morris, "Epidemiology of falls," *Age Ageing,* vol. 30, pp. 3-7, 2001.

[2] WHO. (2007) WHO global report on falls prevention in older age. . *WHO Department of Ageing and Life Course*.

[3] L. Gillespie, W. Gillespie, R. Cumming, S. Lamb, and B. Rowe, "American Geriatrics Society; British Geriatrics Society; American Academy of Orthopaedic Surgeons Panel on Falls Prevention. Guideline for the prevention of falls in older persons Interventions for preventing falls in the elderly," *J Am Geriatr Soc,* vol. 49, pp. 664 - 672, 2001.

[4] K. Berg, "Measuring balance in the elderly: preliminary development of an instrument," *Physiotherapy Canada,* vol. 41, pp. 304-311, 1989.

[5] S. R. Lord, H. B. Menz, and A. Tiedemann, "A Physiological Profile Approach to Falls Risk Assessment and Prevention," *Physical Therapy,* vol. 83, pp. 237-252, March 2003 2003.

[6] M. E. Tinetti, T. Franklin Williams, and R. Mayewski, "Fall risk index for elderly patients based on number of chronic disabilities," *The American Journal of Medicine,* vol. 80, pp. 429-434, 1986.

[7] S. E. Lamb, E. C. Jørstad-Stein, K. Hauer, C. Becker, E. on behalf of the Prevention of Falls Network, and G. Outcomes Consensus, "Development of a Common Outcome Data Set for Fall Injury Prevention Trials: The Prevention of Falls Network Europe Consensus," *Journal of the American Geriatrics Society,* vol. 53, pp. 1618-1622, 2005.

[8] 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, pp. 2918-26, 2010.

[9] M. R. Narayanan, S. J. Redmond, M. E. Scalzi, S. R. Lord, B. G. Celler, and N. H. Lovell, "Longitudinal Falls-Risk Estimation Using Triaxial Accelerometry," *IEEE Trans. Biomed. Eng.,* vol. 57, pp. 534-541, 2010.

[10] M. Marschollek, A. Rehwald, K. H. Wolf, M. Gietzelt, G. Nemitz, H. Meyer Zu Schwabedissen, and R. Haux, "Sensor-based fall risk assessment--an expert 'to go'," *Methods Inf Med,* vol. 50, pp. 420-6, 2011.

[11] D. Giansanti, "Investigation of fall-risk using a wearable device with accelerometers and rate gyroscopes," *Phys. Meas,* vol. 27, pp. 1081-90, 2006.

[12] B. Najafi, K. Aminian, F. Loew, Y. Blanc, and P. A. Robert, "Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly," *IEEE Trans. Biomed. Eng.,* vol. 49, pp. 843-851, 2002.

[13] 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," *Physiological Measurement,* vol. 32, p. 2003, 2011.

[14] J. Howcroft, J. Kofman, and E. D. Lemaire, "Review of fall risk assessment in geriatric populations using inertial sensors," 2013.

[15] T. Shany, S. J. Redmond, M. R. Narayanan, and N. H. Lovell, "Sensors-Based Wearable Systems for Monitoring of Human Movement and Falls," *Sensors Journal, IEEE,* vol. 12, pp. 658-670, 2012.

[16] B. R. Greene, E. P. Doheny, C. Walsh, C. Cunningham, L. Crosby, and R. A. Kenny, "Evaluation of falls risk in community-dwelling older adults using body-worn sensors " *Gerontology,* vol. 58 pp. 472-80, 2012.

[17] S. Mathias, U. Nayak, and B. Isaacs, "Balance in elderly patients: the "get-up and go" test. ," *Arch. Phys. Med. Rehabil.,* vol. 67, pp. 387-9, 1986.

[18] D. Podsiadlo and S. Richardson, "The timed "Up & Go": a test of basic functional mobility for frail elderly persons," *J Am Geriatr Soc,* vol. 39, pp. 142-148, 1991.

[19] A. Shumway-Cook, S. Brauer, and M. Woollacott, "Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test," *Phys Ther,* vol. 80, pp. 896 - 903, 2000.

[20] R. A. Kenny, L. Z. Rubenstein, F. R. Martin, and M. E. Tinetti, "Guideline for the prevention of falls in older people," *J Am Geriatr Soc,* vol. 49, pp. 664-72, 2001.

[21] M. E. Tinetti, M. Speechley, and S. F. Ginter, "Risk factors for falls among elderly persons living in the community," *The New England journal of medicine,* vol. 319, pp. 1701-7, Dec 29 1988.

[22] "Summary of the Updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons," *J Am Geriatr Soc,* vol. 59, pp. 148-57, Jan 2011.

[23] F. Ferraris, U. Grimaldi, and M. Parvis, "Procedure for Effortless In-Field Calibration of Three-Axis Rate Gyros and Accelerometers," *Sensors Mater,* vol. 7, pp. 311-330, 1995.

[24] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning*, 2nd ed.: Springer, 2009.

[25] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence,* vol. 97, pp. 273-324, 1997.