



An automated, electronic assessment tool can accurately classify older adult postural stability



Liam Johnson^{a,b,c}, Adam Fry^d, Behdad Dehbandi^{c,e}, Lawrence Rubin^d, Michael Halem^d, Alexandre Barachant^f, Anna H. Smeragliuolo^{c,e}, David Putrino^{c,d,*}

^aThe Florey Institute of Neuroscience and Mental Health, University of Melbourne, Melbourne, Australia

^bSchool of Health Sciences, Australian Catholic University, Melbourne, Australia

^cDepartment of Telemedicine and Virtual Rehabilitation, Burke Medical Research Institute, White Plains, NY, USA

^dDepartment of Rehabilitation and Human Performance, Icahn School of Medicine, Mount Sinai, New York, USA

^eDepartment of Rehabilitation Medicine, Weill-Cornell Medical College, New York, NY, USA

^fClinical Laboratory for Early Brain Injury Recovery, Burke Medical Research Institute, White Plains, NY, USA

ARTICLE INFO

Article history:

Accepted 1 June 2019

Keywords:

Postural instability
Older adults
Microsoft Kinect 2
Berg Balance Scale

ABSTRACT

Current methods of balance assessment in the clinical environment are often subjective, time-consuming and lack clinical relevance for non-ambulatory older adults. The objective of this study was to develop a novel method of balance assessment that utilizes data collected using the Microsoft Kinect 2 to create a Berg Balance Scale score, which is completely determined by statistical methods rather than by human evaluators. 74 older adults, both healthy and balance impaired, were recruited for this trial. All participants completed the Berg Balance Scale (BBS) which was scored independently by trained physical therapists. Participants then completed the items of the “Modified Berg Balance Scale” in front of the Microsoft Kinect camera. Kinematic data collected during this measurement was used to train a feed-forward neural network that was used to assign a Berg Balance Scale score. The neural network model estimated the clinician-assigned BBS score to within a median of 0.93 points for the participants in our sample population (range: 0.02–5.69). Using low-cost depth sensing camera technology and a clinical protocol that takes less than 5 min to complete in both ambulatory and non-ambulatory older adults, the method outlined in this manuscript can accurately predict a participant’s BBS score and thereby identify whether they are deemed a high fall risk or not. If implemented correctly, this could enable fall prevention services to be deployed in a timely fashion using low-cost, accessible technology, resulting in improved safety of older adults.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

In 2012, older adults in the US were treated for an estimated 1.76 million falls in the emergency room (Burns et al., 2016). Falls are the leading cause of fatal injury, can increase hospital stays and readmissions, and are an important factor affecting morbidity, mortality, and general independence among older adults (Perell et al., 2001). The cost of falls in older adults in the US is high. In 2015, the direct medical costs of fatal falls for people aged 65 years and over totalled \$637.5 million (USD), and \$31.3 billion (USD) was spent on injuries following non-fatal falls (Burns et al., 2016).

Given the risk and cost of falls and fall-related injuries, the assessment of balance and identification of fall risk should be a critical part of routine clinical care for older adults (Xu et al., 2018). Indeed, risk assessment and stratification initiatives may be able to decrease the incidence of falls in older adults, but often require clinical visits that are costly and logistically difficult for both clinicians and patients.

The Berg Balance Scale (BBS) is a commonly used measure of balance within the clinical setting (Berg et al., 1995). However, the BBS has demonstrated floor and ceiling effects in the past (Blum and Korner-Bitensky, 2008), and has limited application in non-ambulatory individuals given most components of the test require the individual to be standing. We previously proposed the modified BBS (MBBS) as an assessment of seated balance for non-ambulatory people, which may enable balance to be safely assessed without the need for supervision (Dehbandi et al.,

* Corresponding author at: Abilities Research Center, 5 East 98th Street, Sub-Basement Room 18, New York, NY 10029, USA.

E-mail address: david.putrino@mountsinai.org (D. Putrino).

2017). The BBS is also a time-consuming assessment (15–20 min), whereas the MBBS can be completed in 5–6 min, which may be particularly advantageous given clinical consults are time limited.

Computerised posturography is a more sensitive measure but typically requires expensive equipment (e.g. force platforms) and trained staff to administer the test and interpret the data collected. Furthermore, the set-up, calibration and data collection are often time-consuming, which might limit its clinical applicability. There are several low-cost, automated, electronic devices that are widely available and have the capacity to measure gait and balance, but until recently, their capacity to sensitively and reliably quantify human motor behaviour has not been established. A comprehensive automated kinematic analysis might offer an innovative solution to the current risk landscape that is the measurement and interpretation of balance in the clinical setting.

The Microsoft Kinect 2 (MK2, Microsoft Corporation, Redmond, WA, USA) is a combined high-definition (HD) video camera and active infrared (IR) camera with a depth sensor designed for 3-dimensional (3D) body tracking. It accurately records the movements of the head, trunk and limbs in 3D space by tracking the displacement of 25 inferred anatomical landmarks; or “centroids”. This technology has widespread application in assessing balance and gait control. It has been shown to accurately assess static (Clark et al., 2015) and dynamic (Eltoukhy et al., 2017) balance, and gait (Mentiplay et al., 2015), including multiple components of the timed-up-and-go (Vernon et al., 2015).

When collected under appropriate conditions, data from the MK2 can be paired with machine learning algorithms to sensitively identify clinically relevant information. Our previous work demonstrated that data collected using the MK2 can be used to discriminate between three distinct states of postural stability in 12 healthy adults performing a modified version of the MBBS (Dehbandi et al., 2017). Taken together, previous work with the MK2 suggests that it could offer a novel approach to identifying individuals with impaired balance. In addition, its low-cost, accessibility and customizability make it, and similar technologies, of relevance to the telemedicine and rehabilitation communities.

The aim of this study was to evaluate the capability of the MK2 to predict the BBS score of older adults with varying levels of postural stability from a few simple, seated movements performed in front of the camera. The research makes progress towards a brief, accurate measure of balance that could be applied both within and away from the clinical setting. This could alleviate time-pressure placed on clinicians/therapists and enhance the robustness of the clinical assessment of balance, and thereby reduce the risk of falling.

2. Methods

2.1. Participants

Participants were recruited from the Burke Rehabilitation Hospital and were included in the study if they were over the age of 18 years, were able to safely sit in an upright position unsupported, could understand and follow verbal commands, and were able to provide informed consent. The participants involved in the study were recruited from a pool of older adults from the Burke Rehabilitation Hospital. This group of participants included stroke survivors (<6 months post-stroke), older adults living with a range of chronic conditions and healthy older adults involved in the wraparound social services provided by the hospital (gym, tai chi classes, aerobic exercise classes). Each participant provided written informed consent prior to their enrolment in the study, which was approved by the Burke Rehabilitation Hospital Committee for Human Rights in Research (BRC-509).

2.2. Data collection

All participants completed a BBS assessment, which was conducted by a trained clinician to provide each participant with a BBS score. The clinical assessors were not required to interact with the technology other than to start and stop the data recording at the beginning and end of the data collection period respectively. Some participants completed two BBS assessments as part of their care, and in these instances the average BBS score was used in the analysis. All participants then completed the MBBS; a series of six separate balance tasks, each involving a simple movement performed in a seated position (Dehbandi et al., 2017).

Participant kinematics were captured at 30 Hz using the MK2, which was positioned 2.7 m in front of the participant. These recordings provided the positions of 25 anatomical landmarks in 3D space over time. These landmarks, identified as ‘centroids’, corresponded to the ‘spine base’, ‘mid-spine’, ‘spine top’, ‘neck’, ‘head’, ‘left shoulder’, ‘left elbow’, ‘left wrist’, ‘left hand’, ‘left thumb’, ‘left fingertip’, ‘right shoulder’, ‘right elbow’, ‘right wrist’, ‘right hand’, ‘right thumb’, ‘right fingertip’, ‘left hip’, ‘left knee’, ‘left ankle’, ‘left foot’, ‘right hip’, ‘right knee’, ‘right ankle’ and ‘right foot’.

2.3. Data analysis

The goal of this data analysis was to build a model to discriminate between individuals at risk of falls from those not at risk from falls using only MK2 recordings of the participants performing the MBBS. Previous analyses have shown that the centroids of the spinal axis contribute the most information on the stability of participants performing the MBBS (Dehbandi et al., 2017). Therefore, to minimize the dimensionality of the model, we selected two of these centroids (‘head’ and ‘mid-spine’) as this was the minimum number of centroids required as inputs to our model. The head centroid was chosen because it provided the most motion of the centroids in the spinal axis, and the mid-spine centroid was chosen as it provides the closest approximation of the participant’s center of mass and it is sufficiently distant from the head to provide a good contrast to the head motion data; and thereby contribute the most information to our model. As centroid position in the vertical direction conveys the least information about a seated participant’s balance, we selected the positional data pertaining to the medial-lateral and anterior-posterior directions for use in our model. Finally, a previous regression analysis of the data generated from each of the six tasks in the MBBS determined that tasks five and six generated the most useful information for estimating a participant’s total BBS. We therefore restricted our analysis to data from tasks five and six. The instructions for these tasks were as follows:

Task 5: Pick up an object from the floor with your left hand, and return to sit upright, unsupported. Lean over to place the object back on the floor. Repeat using the right hand. End in an upright seated position.

Task 6: Start with arms crossed, sitting unsupported. Raise one arm 90° out from your side. Hold for 10 s, then return to the start position for 10 s. Repeat with the other arm.

This meant our analysis was limited to time-series data representing two dimensions of physical space, from two centroids within the upper body, and from two of the six tasks of the MBBS.

To estimate total BBS scores from the MK2 data we used a feed-forward neural network model; a type of machine learning algorithm. The configurable model was hand-written in C# and used back propagation of errors with a steepest descent learning method. Inputs to the model included the variance/covariance of each time-series as determined by the unadjusted root mean square standard deviation (σ) of the centroid position data ($\sigma_{\text{Head}_X}^2$, $\sigma_{\text{Head}_Z}^2$, $\sigma_{\text{Mid-spine}_X}^2$, $\sigma_{\text{Mid-spine}_Z}^2$, $\sigma_{\text{Head}_X\text{-Head}_Z}^2$, $\sigma_{\text{Head}_X\text{-Mid-spine}_X}^2$,

$\sigma_{\text{Head}_X\text{-Mid-spine}_Z}^2$, $\sigma_{\text{Head}_Z\text{-Mid-spine}_X}^2$, $\sigma_{\text{Head}_Z\text{-Mid-spine}_Z}^2$, and $\sigma_{\text{Mid-spine}_X\text{-Mid-spine}_Z}^2$; where x = medial-lateral direction, y = superior-inferior direction and z = anterior-posterior direction) for MBBS tasks five and six for each participant. In addition, the BBS scores were included as the variable to be estimated from the centroid position data. The total number of input nodes in the neural network was therefore 21. The model also included two layers of 11 hidden nodes per layer, and one output node corresponding to the estimated total BBS scores.

3. Results

A total of 74 participants (35 female) were recruited for this study. Participants were primarily older adults with an average age of 73.1 years (46–95). Participants included 59 individuals with chronic disease (i.e. cardiovascular disease, musculoskeletal injury, neurological conditions), of which 23 were experiencing multi-morbidity. In addition to this, 15 otherwise healthy older adults were recruited into the study. No adverse events were reported.

In order to perform our intended analysis, it was necessary to sort our data into “training” (data that is used to develop a classification algorithm) and “testing” (data that is used test the accuracy of the classification algorithm developed with the training set) datasets. We collected data from 74 study participants, and our initial approach was to create training and testing datasets by sorting participants by their BBS score, and then assigning odd numbered participants to the training dataset and even numbered participants to the testing dataset. However, due to software errors that occurred during data collection that were only discovered post-hoc, three participants were found to have incomplete datasets, and were thus considered ineligible to be part of the training dataset. Furthermore, a linear regression model approach determined that five additional participants were outliers due to variance in Kinect centroid motion during some of the tasks that was two standard deviations or greater than the population datasets. This sort of variation likely occurred due to poor skeletal tracking performance of the Kinect camera during task performance in these participants. Inclusion of these outliers in the Training dataset would have reduced the fit of the model to the training dataset. They were subsequently excluded from the training dataset. Thus, data from 31 participants were included in the Training set and data from the remaining 43 participants (including those with incomplete data and those identified as outliers) were then included in the Testing dataset.

The training dataset included participants with BBS scores between 18 and 56, while test dataset included participants with BBS scores between 21 and 56, with no participants common to both datasets. Using the training dataset to populate the input nodes, the neural network was initiated in a random state and continued training until the mean squared error of the output (estimated BBS scores) was less than a threshold of 1×10^{-6} , which corresponded to approximately 1.5 million iterations. The resultant model weightings were recorded, and this iterative process was repeated a total of 100 times. The mean model weightings across these 100 repetitions were then used to predict the BBS scores of the test dataset.

The results of the neural network analysis are shown in Table 1. These results reveal that the neural network model outlined in the methods was able to estimate the clinician assigned BBS score to within a median of 0.93 points for the participants in our sample population (range: 0.02–5.69). Moreover, using a threshold of <40 to stratify participants as either high fall risk or moderate-low fall risk (Shumway-Cook et al., 1997), our model was able to correctly classify all except for one of the participants in the test

Table 1
The clinician assigned BBS scores and the estimated BBS scores from the neural network model, for each of the 43 participants in the test dataset.

Clinician BBS score	Estimated BBS score	Absolute error
21	19.06	1.94
28	24.56	3.44
33	27.31	5.69
36	34.82	1.18
38	37.83	0.17
39	37.91	1.09
40	38.52	1.48
42	41.32	0.68
42	39.95	2.05
42	42.52	0.52
43	42.21	0.79
44	41.17	2.83
45	46.21	1.21
47	50.27	3.27
48	47.82	0.18
48	47.55	0.45
48.5	48.52	0.02
49	49.80	0.80
50	51.86	1.86
50	50.33	0.33
51	50.21	0.79
51	49.53	1.47
51	52.08	1.08
52	52.14	0.14
52	52.93	0.93
52	53.03	1.03
52.5	52.95	0.45
53	54.16	1.16
53	51.85	1.15
53	52.69	0.31
53	53.18	0.18
53.5	53.76	0.26
54	53.64	0.36
54	52.88	1.12
54	54.41	0.41
55	53.77	1.23
55	54.21	0.79
55	54.40	0.60
55.5	54.60	0.90
55.5	54.61	0.89
56	53.08	2.92
56	55.00	1.00
56	54.94	1.06
	Median (IQR):	0.93 (0.45–1.23)

BBS, Berg Balance Scale; IQR, Interquartile Range.

dataset, with the one misclassified participant receiving a score of exactly 40 from the clinician assessment.

4. Discussion

In this study, the MK2, a low-cost depth sensing video camera, was paired with sophisticated modelling to develop a method of stratifying fall risk in older adults. Our model was highly accurate in predicting participants as either high fall risk (i.e. <40) or moderate-low fall risk (i.e. ≥40), with only one of 43 participants misclassified in our test dataset. Moreover, the median error of our model’s predictions was less than one point on the BBS. This protocol may therefore hold substantial clinical value given it is a safe and time-efficient balance assessment that can be performed by both ambulatory and non-ambulatory individuals.

The machine learning algorithm accurately predicted the BBS score from the MK2 recordings of two simple, seated balance tasks. This approach offers several advantages over a standard BBS assessment. For example, each task in our MBBS takes less than 2 min to complete, making our proposed method of balance assessment significantly faster to perform than the standard BBS (i.e. 2–4 min vs. 15–20 min for the complete BBS). Furthermore, because

all the tasks required in the MBBS are performed relatively slowly and in a seated position, the risk of falling during an assessment is much lower than the BBS.

There have been few comparisons between remote versus in-person ratings of the BBS. Whilst the standard BBS has been shown to have acceptable responsiveness to change in chronic stroke (Alghadir et al., 2018), the median error in the predicted BBS scores (<1 point) was considerably lower than the minimal detectable change in stroke survivors (5.97 points at 80% confidence level) (Saso et al., 2016). This suggests the methods proposed in this study may be more capable of tracking longitudinal changes in fall risk with enhanced sensitivity compared with common clinical practice methods. A previous challenge for tele-assessment via video capture was the need for repeated task performance to capture both frontal and lateral views (Venkataraman et al., 2017). The utility of the MK2 to provide high definition 3D body tracking enables our method to overcome this problem. Furthermore, whilst functional balance measures such as the Performance Oriented Mobility Assessment (POMA) (Tinetti, 1986) include a seated balance component, this makes only a small contribution to the overall measure.

The MK2 has previously been shown to provide reliable and valid gait assessments in both healthy adults and stroke survivors (Clark et al., 2015; Mentiplay et al., 2015). Our previous study, which benchmarked the validity of the MK2 against force platforms, demonstrated the MK2 could be used to measure postural stability in healthy adults (Dehbandi et al., 2017). The current study advances this work by demonstrating the accuracy of the MK2 to estimate postural stability in a cohort of older adults, including some with significant balance issues. The excellent predictive capacity of this model demonstrates progress towards a tool that can automatically and accurately assess an individual's balance ability. The methodology presented in this study offers the potential for remote monitoring of fall-risk from the safety of the patient's home. The equipment used in this study is inexpensive, widely available and unobtrusive; improving access to the monitoring of fall-risk in clinical populations experiencing postural instability. We acknowledge that the methodological requirements (i.e. appropriate space, access to mains power via cables), and the potential need for assistance with the initial set-up, may limit the practicality of our assessment technique. Research into the acceptability and usability of depth-sensing devices and skeletal tracking for physical function assessments in the home environment is needed.

Although the MK2 captures full body motion across 25 different centroids, in our predictive model, we limited ourselves to the use of two centroids - the lowest number of centroids appropriate for our chosen machine learning methodology. We are interested in the lowest number of centroids to provide accurate estimations of fall-risk for two main reasons: (1) It allows for this otherwise computationally demanding analysis to be performed in real time if required, and (2) it makes our methodology generalizable in the future to much simpler methods of motion capture techniques and devices, such as wearable sensors.

This study is not without its limitations. We acknowledge the low number of participants assessed who might be considered as a high fall-risk (i.e. <40) may limit the generalizability of our findings. Future trials will aim to reproduce this finding with a larger sample size that also includes a wider spectrum of balance ability. As soon as our findings are confirmed in a larger population, steps will be taken to automate our machine learning approach to BBS classification so that this system can be deployed as a truly "plug-and-play" system for untrained clinical users. This will help us to address any potential overfitting issues that may have occurred from having a smaller sample size. Furthermore, all participants included in this study were ambulatory, and it is possible

non-ambulatory older adults respond differently to the MBBS protocol used in this study. This would be an important future direction of this work with substantial clinical relevance. Many people with conditions such as spinal cord injury have an extremely high risk of falling, but often don't receive balance assessments because few clinical scales exist to stratify fall-risk in non-ambulatory individuals. Although the MK2 was discontinued by Microsoft shortly after the completion of this study, products that are similar in price and quality, such as the Orbbec Astra, can be used for the data collection component of this protocol. In addition, Microsoft has recently announced the release of the Microsoft Kinect 3 in 2019, bringing the possibility of even more accurate technology. A recent narrative review highlights the current developments in 3D camera systems and skeleton tracking software, including alternatives to the now defunct MK2 (Clark et al., 2019).

5. Conclusion

We have shown that we can capture and interpret kinematic data to generate accurate predictions of balance scores in older adults with and without balance impairments. This methodology has the potential to support clinicians and therapists in their clinical decision-making when assessing the balance and potential fall-risk of their patients. It is also highly relevant to telerehabilitation service providers as we now have the potential to accurately measure balance, and the risk of falling, in a home setting using low-cost, accessible technology. The practicality of a patient, and/or their carer, setting up and using a device like the MK2 within their home needs exploration.

Acknowledgements

We wish to acknowledge Dr John Long, Mr Silverio Bumanlag, Mr Victor He and Ms Anna Lampe for their contribution to the study. We would also like to acknowledge the study participants who gave up their time to participate in this study.

Conflict of interest

The authors have no conflicts.

Sponsor's role

No sponsors were involved in this study.

References

- Alghadir, A.H., Al-Eisa, E.S., Anwer, S., Sarkar, B., 2018. Reliability, validity, and responsiveness of three scales for measuring balance in patients with chronic stroke. *BMC Neurol.* 18, 141.
- Berg, K., Wood-Dauphinee, S., Williams, J.L., 1995. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scand. J. Rehabil. Med.* 27, 27–36.
- Blum, L., Korner-Bitensky, N., 2008. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys. Ther.* 88, 559–566.
- Burns, E.R., Stevens, J.A., Lee, R., 2016. The direct costs of fatal and non-fatal falls among older adults – United States. *J. Safety Res.* 58, 99–103.
- Clark, R.A., Mentiplay, B.F., Hough, E., Pua, Y.H., 2019. Three-dimensional cameras and skeleton pose tracking for physical function assessment: a review of uses, validity, current developments and Kinect alternatives. *Gait Posture* 68, 193–200.
- Clark, R.A., Pua, Y.H., Oliveira, C.C., Bower, K.J., Thilarajah, S., McGaw, R., Hasanki, K., Mentiplay, B.F., 2015. Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control. *Gait Posture* 41, 210–213.
- Dehbandi, D., Barachant, A., Smeragliuolo, A.H., Long, J.D., Bumanlag, S.J., He, V., Lampe, A., Putrino, D., 2017. Using data from the Microsoft Kinect 2 to determine postural stability in healthy subjects: a feasibility trial. *PLoS One* 12, e0170890.

- Eltoukhy, M., Kuenze, C., Oh, J., Wooten, S., Signorile, J., 2017. Kinect-based assessment of lower limb kinematics and dynamic postural control during the star excursion balance test. *Gait Posture* 58, 421–427.
- Mentiplay, B.F., Perraton, L.G., Bower, K.J., Pua, Y.-H., McGaw, R., Heywood, S., Clark, R.A., 2015. Gait assessment using the Microsoft Xbox One Kinect: concurrent validity and inter-day reliability of spatiotemporal and kinematic variables. *J. Biomech.* 48, 2166–2170.
- Perell, K.L., Nelson, A., Goldman, R.L., Luther, S.L., Prieto-Lewis, N., Rubenstein, L.Z., 2001. Fall risk assessment measures: an analytic review. *J. Gerontol. A Biol. Sci. Med. Sci.* 56, M761–766.
- Saso, A., Moe-Nilssen, R., Gunnes, M., Askim, T., 2016. Responsiveness of the Berg Balance Scale in patients early after stroke. *Physiother. Theory Pract.* 32, 251–261.
- Shumway-Cook, A., Baldwin, M., Polissar, N.L., Gruber, W., 1997. Predicting the probability for falls in community-dwelling older adults. *Phys. Ther.* 77, 812–819.
- Tinetti, M.E., 1986. Performance-oriented assessment of mobility problems in elderly patients. *J. Am. Geriatr. Soc.* 34, 119–126.
- Venkataraman, K., Morgan, M., Amis, K.A., Landerman, L.R., Koh, G.C., Caves, K., Hoenig, H., 2017. Tele-assessment of the Berg Balance Scale: effects of transmission characteristics. *Arch. Phys. Med. Rehabil.* 98, 659–664.
- Vernon, S., Paterson, K., Bower, K., McGinley, J., Miller, K., Pua, Y.H., Clark, R.A., 2015. Quantifying individual components of the timed up and go using the Kinect in people living with stroke. *Neurorehabil. Neural Repair.* 29, 48–53.
- Xu, T., Clemson, L., O'Loughlin, K., Lannin, N.A., Dean, C., Koh, G., 2018. Risk factors for falls in community stroke survivors: a systematic review and meta-analysis. *Arch. Phys. Med. Rehabil.* 99, 563–573.