Article

The Validity of a Mixed Reality-based Automated Functional Mobility Assessment

Ruopeng Sun [1,] *, Roberto G. Aldunate [1] and Jacob J Sosnoff [1]

1 Department of Kinesiology and Community Health, University of Illinois at Urbana-Champaign; rusun@illinois.edu, aldunate@illinois.edu, jsosnoff@illinois.edu;

* Correspondence: rusun@illinois.edu;

Abstract: Functional mobility assessments (i.e. Timed Up and Go) are commonly used clinical tools for mobility and fall risk screening in the aging population. In this work, we proposed a new Mixed Reality (MR)-based assessment that utilized a Microsoft HoloLens™ headset to automatically lead and track the performance of functional mobility tests, and subsequently evaluated its validity in comparison with reference inertial sensors. Twenty-two healthy adults (10 older, 12 young) participated in this study. An automated functional mobility assessment app was developed based on the HoloLens platform. Mobility performance was recorded with the headset built-in sensor and validated with reference inertial sensor (Opal, APDM) taped on the headset and lower back. Results indicate vertical kinematic measures by HoloLens was in good agreement with the reference sensor (Normalized RMSE ~ 10%). Additionally, the HoloLens-based test completion time was in perfect agreement with clinical standard stopwatch measure. Overall, our preliminary investigation indicates that it is possible to use an MR headset to automatically guide users to complete common mobility tests with good measurement accuracy, thus it has great potential to provide objective and efficient sensor-based mobility assessment.

Keywords: Mixed Reality Headset; Mobility Assessment; Wearable Sensor; Fall Risk; Aging

---

1. Introduction

Falls are the leading cause of injury related death in older adults [1]. Over 1 in 4 older adults will experience a fall in the next year and a significant portion of those that fall will suffer an injury [2], resulting in more than $50 billion annual medical costs [3]. Moreover, although most falls do not end in death or result in significant physical injury, the psychological impact of a fall, such as fear of falling, loss of confidence, often results in anxiety and depression that further decreased quality of life [4]. The risk of falling increase with aging due to multiple risk factors, such as deficits in vision, cognition, muscle strength and mobility [1]. Given the frequency and severe consequences of falls, there is a critical need for early and regular monitoring on individual's fall risk to reduce falls and fall-related injuries.

Indeed, the American Geriatrics Society (AGS) and the Centers for Disease Control and Prevention (CDC) recommends annual fall risk screening for older adults [5, 6]. The most commonly used fall risk screening tests include functional mobility tests such as the five time sit to stand test (STS) [7] and timed up and go test (TUG) [8]. Both tests are valid and reliable clinical tests focusing on assessing lower-limb muscle strength and mobility, which are both critical factors to the overall risk of falling. The STS is a clinical test that asks participants to stand up and sit down from a chair five times as quickly as possible without using the armrest. Whereas the TUG test requires participant stand up from a chair, walks 3m at normal pace, turn around and return to the chair.

Even though such standardized tests are relatively easy to conduct, it is still underutilized and not routinely integrated into clinical practice. Partially due to clinicians' time constraints and competing medical priorities, lack of access to lab-grade advanced testing equipment (such as motion capture device, force platform), as well as lack of clinical expertise [9], thus limiting access to fall risk screening in the community dwelling older adults. Consequently, older adults remain unaware of their individual fall risk, appropriate fall prevention approaches and at elevated risk of falls.

With the recent advancement in sensing technology, sensor-based fall risk assessment that can efficiently capture and analyze quantitative mobility data have received a growing interest for its portability, accessibility and inexpensiveness [9, 10]. More specifically, the use of wearable sensors (inertial measurement unit, IMU) for mobility-related tracking has been the focus of these work, in which miniature accelerometers and/or gyroscopes were used to quantify movement pattern/abnormality by various time and frequency parameters [9, 11].

Two recent systematic reviews on sensor-based fall risk assessment [9, 12] identified over 50 investigations using IMUs for fall risk assessment in older adults. This body of literature highlights that sensors are capable of accurately quantifying mobility and capable of distinctive in identifying mobility impairment in high fall risk individuals. Overall it is concluded that wearable sensors are a viable technology for fall risk assessment. It is worth noting that in most investigations, IMU sensor(s) was attached to lower back and/or lower limb as stand-alone recording device, and still required additional personnel to guide the wearer through the assessment protocols.

An alternative approach to further improve the efficiency of sensor-based assessment is to offer direct technology interaction with the intended users, such as deliver demonstrations and instructions, and receive user inputs. Such a system would be able to provide an automated and self-guided assessment system that requires no additional personnel. To achieve this goal, an ideal device should be able to communicate with the wearer with both visual and auditory prompt, and allow user input through natural interactions (gesture, voice, gaze, etc.). Indeed, recent research has highlighted that smartphone and tablets can provide valid and reliable fall risk assessment to older adults [13, 14].

Mixed reality head-mounted display (HMD, e.g. Microsoft HoloLens, Fig.1a) systems are also uniquely fitted for technology-based fall risk assessment. For instance, the HoloLens uses a transparent display with light projector to provide holograms on the lenses in front of user's eye that blend the digital display with physical environments [15]. It contains multiple sensors to scan the user's environment which enables the holograms to be placed at a specific location in real world [15]. By using such device in mobility assessment, the user can receive instructions and visual demonstration, naturally interact with the virtual display through voice command, gesture control and gaze, and complete mobility tests with full visibility of the surrounding environments. In addition, the embedded IMU and depth sensor can be used to track the user's head movement during balance and mobility tests. Although less commonly used than sensors mounted at lower limb and lower back for mobility assessment, head movement has been used as approach for mobility evaluation, based on the notion that head movement is linked to the trunk movement as well as gait-related oscillations during locomotion [16, 17]. Additionally, given that head stabilization has been shown as a critical component in maintaining upright posture [16], monitoring head movement may provide novel insights to the mobility control and fall risk evaluation. These unique features of HMD have the potential to enable older adults to complete fall risk screening intuitively and autonomously.

Although this Mixed Reality Headset holds promise for enhancing the fall risk assessment in community dwelling older adult, its validity for objective mobility assessment have not been investigated. Therefore, the aim of this study is to evaluate the validity of the mixed-reality headset for automated mobility assessment in young and older adults.

2. Materials and Methods

2.1 Participants

Twenty-two healthy adults (10 older adults -OA, 12 young adults-YA) participated in this study. The inclusion criteria for participation were age between 18-30 years or 65 years and older, able to stand 30s unaided, able to walk with or without aid, have normal or corrected to normal hearing and vision, no history of neuro-muscular or cardiovascular disease, and no history of motion sickness, chronic neck pain or seizure-related conditions. All procedures were approved by the University of Illinois at Urbana-Champaign Institutional Review Board, and all participants completed written informed consent prior to participation. Participant testing was performed at 2 sites selected for convenience for the participants. OA were tested at an unoccupied apartment setting at a local retirement community, while YA were tested at the university laboratory setting.

2.2 System Setup

A customize built Universal Windows Platform application was developed under Unity (2018 2.6 personal) and Visual Studio (Microsoft Visual Studio 2017), and deployed on the Microsoft HoloLens head-mounted display operating under the Windows 10 system. The HoloLens features depth cameras for environment scanning and spatial mapping, as well as IMU for position and orientation estimation [15]. The transparent visor and light projector allow user to see high-definition virtual content (hologram) over real world objects (Fig.1b). The field of view from HoloLens was estimated as 30° H and 17.5° V [15]. The system can operate as a stand-alone device that requires neither PC nor smartphone. For this project, the onscreen display was also streamed on a laptop for monitoring participant's interaction with the system. The HoloLens features multimodal user interaction methods, such as finger pinch, voice command, and estimated gaze fixation, etc. In order to simplify user interaction and allow intuitive control for senior user, we chose to use gaze fixation (orientation estimated) to control the interface, i.e. user will control the system by fixating their gaze on a control button for 1-2 s (Fig 1b, purple circle as gaze cursor, white blocks as control buttons. Video in Supplementary S1). Additionally, to facilitate user onboarding and ensure self-guided test completion, participant watched a standardized tutorial video on a laptop explaining how to put on, adjust and control the headset (Supplementary S2). Participants were encouraged to ask questions before putting on the headset. Details about the user-interface design process and usability investigation will be reported in a separate publication.

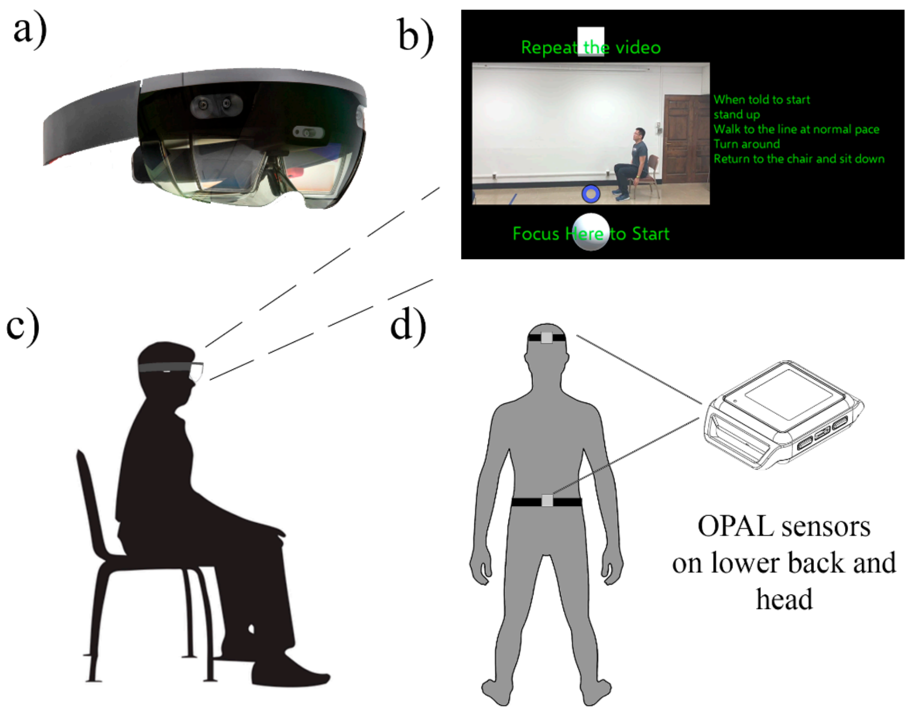

Figure 1. Illustration of system setup. (**a**) HoloLens headset; (**b**) Onscreen hologram instruction for Timed Up and Go test (Shaded background, video animation, green font, white control button and purple gaze cursor); (c) Illustration of participant's starting position; (d) Reference IMU sensors placement.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2019 doi:10.20944/preprints201904.0120.v1

Peer-reviewed version available at Sensors 2019, 19, 2183; doi:10.3390/s19092183

Based on the CDC fall risk assessment recommendations (Stopping Elderly Accidents, Death & Injuries - STEADI [6]) and feasibility of head-mount movement tracking, a set of valid and reliable clinical tests focusing on mobility and muscle strength were selected and integrated in the automated mobility assessment App. Muscle strength and coordination was assessed with the STS test, whereas mobility was assessed with TUG test.

2.3 Test Procedure

Participants completed an MR-based mobility assessment as well as a clinical fall risk assessment. During the MR-based assessment, participants were outfitted with two additional APDM Opal IMU sensors (APDM, Inc.). One was secured to the top of the HoloLens visor – denoted as HD sensor and one placed on participant's lower back via belt – denoted as LB sensor, Fig 1d). After being fitted with the headset, participants were prompted to complete test sessions in the following order: 1) STS, 2) TUG. For each test, a video recording with standard demonstration and instruction was displayed on the headset (Fig. 1b, center display and text instruction). Upon video completion, participants were provided the option to proceed to test or repeat the demonstration (Fig. 1b, white blocks). During the test, a 5s count down with audio tone and text of the instruction was displayed. Test completion button was prompted up after a 15s delay. After test completion, participants had options to repeat the test if not satisfied with their performance. To ensure identical test setup between participants, markings were placed on the ground (standing feet placement, chair location and 3 m walking path). Research personnel offered safety spotting and minimal interaction with the participant, unless asked by the participant to help. Participant who made errors during test (i.e. false start, incorrect number of repetitions, etc.) were asked to repeat the test.

After completion of the self-guided mobility assessment, the physiological profile assessment (PPA) [18] was administered by trained research personnel to evaluate the overall fall risk. The Montreal Cognitive Assessment (MoCA) [19] and the Activity-specific Balance Confidence (ABC) scale [20] were also administered to assess participant's cognitive function and balance confidence, respectively. The MoCA test is a validated screening tool for detecting cognitive impairment, whereas the ABC scale is a validated self-reported questionnaire of confidence in performing various daily activities without losing balance. The PPA consists of a set of comprehensive tests assessing vision, lower limb sensation, muscle strength, reaction time and balance that are associated with risk of falling [18].

2.4 Data Processing

Due to HoloLens API setup on sensor data access, raw accelerometry/gyroscope data was not accessible, thus only the processed head position and orientation was available for recording at a dynamic sampling rate at/around 30Hz (variation due to windows internal clock frequency). Such data was processed by its internal proprietary sensor fusion algorithm (IMU, depth and environmental cameras) that output the 3D head spatial coordinate and gravity aligned orientation. Acceleration and gyroscope data from the Opal sensor was recorded at 128Hz, and gravity corrected after orientation estimation using an extended Kalman filter provided by APDM. Both HoloLens and Opal data was segmented and synced to each task, resampled at 30 Hz and low-pass filtered (4th order Butterworth) with cutoff frequency at 5 Hz [21-23].

For STS and TUG tests, the gravity corrected vertical (VT) acceleration data from Opal sensors were double integrated over time to obtain the vertical displacement, with drift and integration error corrected using 1) a high pass filter (4th order Butterworth, 0.1Hz) [23] and 2) drift correction under the assumption that participants reach the same height when they make contact with the chair (zero displacement update-ZDU) [24]. The processed VT displacement from HoloLens and Opal sensors was then time-aligned using cross correlation analysis (calculating the similarity and time lag between signal). Finally, time-aligned VT kinematic data (displacement, velocity, and acceleration) profile were derived from Opal and HoloLens using numerical integration and differentiation accordingly. The VT data were also used to calculate the following performance features: 1) STS duration: time between the initiation of first chair rising and the completion of last

Peer-reviewed version available at Sensors 2019, 19, 2183; doi:10.3390/s19092183

chair descend [21, 22]. 2) STS mean duration of the sitting and standing phase [21, 22]. 3) Maximum acceleration and velocity during STS. 4) TUG duration: time between the initiation of chair rising and the completion of chair descend. 5) Maximum acceleration and velocity during TUG. Due to significant signal drift over time in AP, ML direction using Opal sensor (see discussion), and lack of viable signal correction method, signal comparison in AP/ML direction was not performed for STS and TUG tests.

Pairwise signal agreement between HoloLens, HD and LB sensors were analyzed using the normalized root mean squared error (NRMSE – RMSE divided by signal amplitude range), as well as the cross-correlation coefficient (correlation coefficient at zero lag, denoted as Xcor). The NRMSE reports the error as a ratio of the measurement range, with lower values indicates better signal agreement [25]. The Xcor measure the signal similarity of two time series, with higher value (~1) indicates better signal agreement [23]. Mean and 95% confidence interval of NRMSE and Xcor were calculated by the functional test condition (STS, TUG).

STS and TUG duration derived from HoloLens measure was compared with manual stopwatch tracking using Bland-Altman limit of agreement analysis [26]. The Bland-Altman limit of agreement is a robust statistical approach to indicate the level of agreement between any two measurements. Since a high correlation between any two methods does not necessarily mean that the two methods are in good agreement, the Bland–Altman technique is utilized in many studies to investigate the presence of absolute agreement between the two technologies.

For all derived mobility features from the HoloLens (STS duration, STS mean duration of the sitting and standing phase, Maximum acceleration and velocity during STS, TUG duration, Maximum acceleration and velocity during TUG), group (OA and YA) comparison was also conducted using two-tail student t-test. All data processing and statistical analysis was performed with customized MATLAB program (MathWorks, Inc.)

3. Results

Table 1. Participant characteristics (mean and standard deviation). * indicates significant group difference ($p < 0.05$)

	OA n=8,6F	YA n=12, 6F
Age (yrs) *	78.2 (6.1)	24.4 (3.9)
BMI (kg/m^2)	23.9 (3.6)	24.5 (2.9)
MoCA *	26.2 (2.3)	28.6 (1.7)
ABC	88.8 (13.3)	96.0 (3.7)
MET *	19.9 (1.5)	21.2 (0.6)
RT (ms) *	257.7 (33.6)	217.5 (32.8)
Proprio	3.0 (1.2)	3.3 (3.5)
KneeMax (kgf) *	25.3 (9.6)	41.9 (8.5)
AP sway (mm)	27.2 (9.8)	20.8 (10.9)
ML sway (mm)	33.7 (18.9)	20.5 (12.3)
PPA *	0.9 (0.7)	-0.3 (0.7)

Two older participants (82 and 91 years old, both male and have significant fall risk -PPA >2) did not complete the self-guided assessment, due to balance/mobility deficits and need for assistance in challenging conditions. Therefore, only 8 OA and 12 YA were included in data analysis. Additionally, 2 OA participants were asked to repeat the test due to performance error (false start, wrong number of sit to stand repetition, etc.). Participants' demographic characteristics and physiological profile measured by PPA are presented in Table.1. As expected, significant difference in age, cognitive function (MoCA), contrast visual acuity (MET), reaction time, muscle strength, and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2019 doi:10.20944/preprints201904.0120.v1

Peer-reviewed version available at Sensors 2019, 19, 2183; doi:10.3390/s19092183

overall risk of falls were observed between OA and YA. It is worth noting only 1 of the remaining 8 OA has significant fall risk (PPA > 2).

Representative data traces of STS and TUG test from a young participant (27 years old, female) are illustrated in Fig 2. Vertical kinematic data comparison between HoloLens and Opal sensors using NRMSE and Xcor are presented in Table 2. Overall, data recorded from HoloLens (green line in Fig 2) is in good agreement with the HD/LB sensors (red/blue line in Fig 2). In general, signal similarity as measured by the Xcor is good to excellent (0.740-0.998), with higher Xcor observed in displacement measure (0.965-0.998) in comparison to velocity (0.853-0.979) and acceleration (0.740-0.888). This finding can be explained as the displacement drift was corrected using the ZDU method, whereas velocity and acceleration signal from HD/LB sensors remain slightly affected by the integration/differential error. For STS task, signal agreement as measured by NRMSE were excellent (below 10%) for all measures except displacement comparison between HoloLens and LB sensor (11.88%). Whereas for TUG task, signal agreement was relatively low (close to 20% NRMSE for displacement comparison), likely due to ZDU drift correction only performed 1 time over the entire recording for HD/LB sensor (contrary to drift correction after each sitting cycle in STS task), resulting in more vertical displacement bias (Fig 2e).

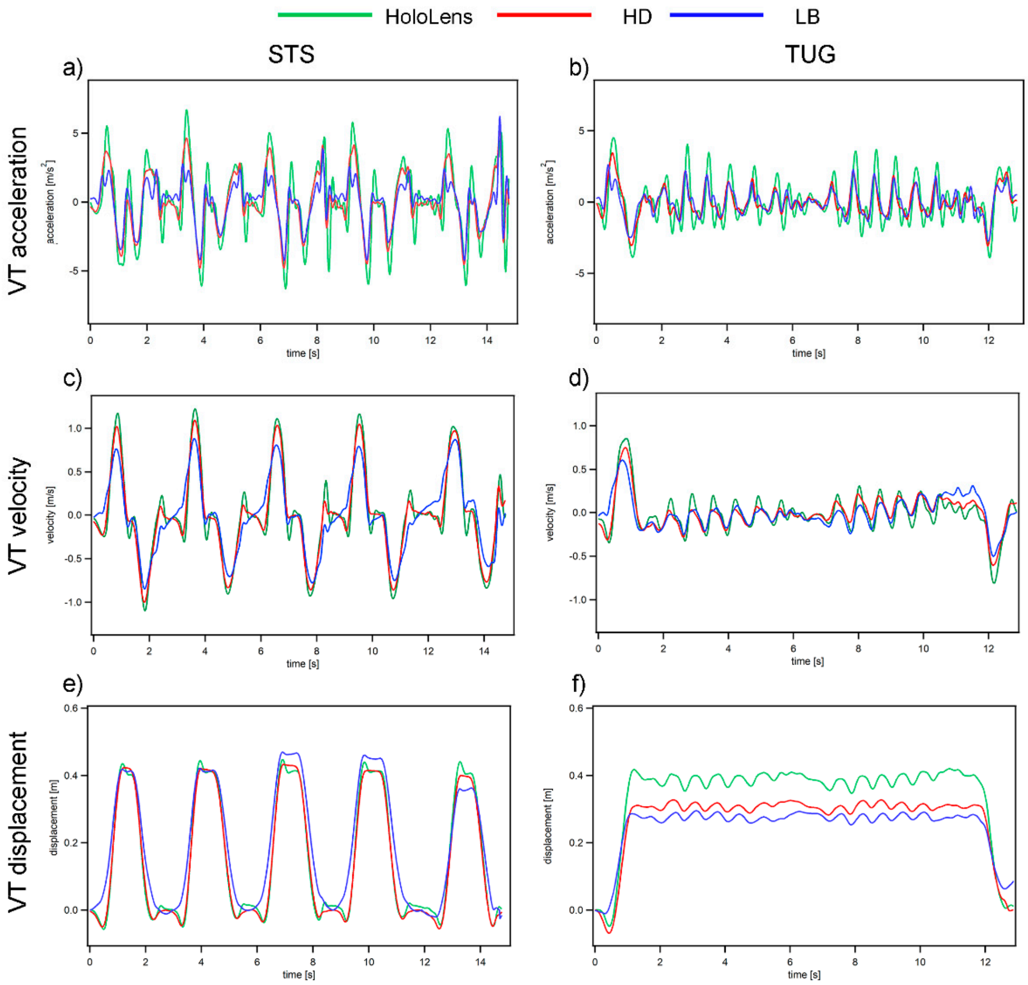

Figure 2. Sample kinematic profile from a young participant. Green denotes HoloLens, Red denotes HD sensor, Blue denotes LB sensor. (**a, b**) VT acceleration profile from STS and TUG task; (**c, d**) VT velocity profile from STS and TUG task; (**e, f**) VT displacement profile from STS and TUG task.

Peer-reviewed version available at *Sensors* **2019**, *19*, 2183; doi:10.3390/s19092183

Table 2. Kinematic measurement (VT acceleration, velocity, and displacement) agreement between HoloLens and HD/LB sensors. NRMSE – Normalized Root Mean Squared Error. Xcor – Cross Correlation Coefficient. All values reported as mean and 95% confidence interval.

HoloLens vs HD				
		A	**V**	**D**
STS	**NRMSE**	9.60 (8.70,10.51)	4.83(4.22,5.45)	5.58 (4.29, 6.87)
	Xcor	0.888(0.872,0.904)	0.979(0.975,0.983)	0.993 (0.989 0.997)
TUG	**NRMSE**	10.53 (9.60,11.46)	6.16 (5.57,6.76)	19.56 (17.24,21.87)
	Xcor	0.802 (0.770,0.834)	0.926(0.918,0.934)	0.998 (0.997 0.999)
HoloLens vs LB				
		A	**V**	**D**
STS	**NRMSE**	9.77 (8.29,11.25)	8.55 (6.98,10.12)	11.88 (9.72,14.03)
	Xcor	0.765 (0.704,0.827)	0.900(0.851,0.949)	0.965 (0.949,0.982)
TUG	**NRMSE**	8.48 (7.56,9.41)	7.68 (7.05,8.31)	14.07 (11.86,16.28)
	Xcor	0.740 (0.695,0.786)	0.853 (0.835,0.872)	0.986 (0.978 0.993)

Figure 3a and 3b shows the Bland-Altman plot for the agreement in STS and TUG completion time between HoloLens and manual stopwatch recording. The absolute difference between each data pair is plotted against their mean. The two horizontal lines represent the 95% limits of agreement (range of error) calculated as 1.96 times the standard deviation from the mean differences between two methods. The figure illustrates that the mean difference between two methods is less than 0.02 s for STS and 0.13 s for TUG measure, with range of error within ± 0.8s.

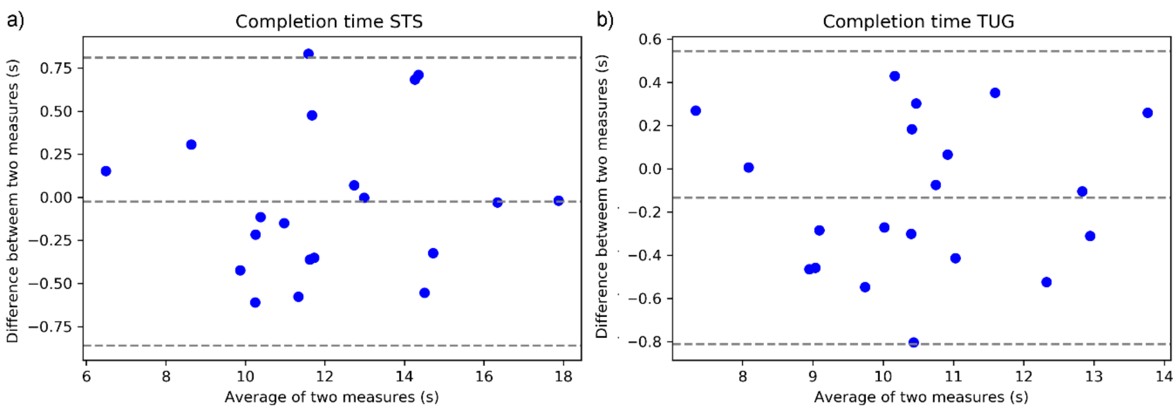

Figure 2. Bland-Altman plot of sensor-derived and stopwatch timed task completion time. (a) STS completion time; (b) TUG completion time.

Peer-reviewed version available at *Sensors* **2019**, *19*, 2183; doi:10.3390/s19092183

Table 3 shows the group comparison (OA vs YA) in functional performance outcomes derived from the HoloLens measures. Overall, there is no significant difference between groups, with marginal significant difference observed in max velocity in STS and TUG task, reflecting the healthy nature of the OA samples (only 1 of 8 participants who completed the assessment have significant fall risk).

Table 3. Group differences of key outcome measures (mean and standard deviation).

Task	Outcome Measures	OA	YA	*p*
STS	Total Time (s)	12.22 (3.61)	12.08 (1.99)	0.922
	Mean Stand Time (s)	0.52 (0.18)	0.64 (0.22)	0.198
	Mean Sitting Time (s)	1.15 (0.55)	1.03 (0.23)	0.575
	Max Acceleration (m/s²)	4.75 (1.81)	6.22 (2.03)	0.108
	Max Velocity (m/s)	1.02 (0.16)	1.20 (0.28)	0.087
TUG	Total Time (s)	10.61 (2.37)	10.56 (1.00)	0.96
	Max Acceleration (m/s²)	3.98 (0.92)	3.97 (0.62)	0.961
	Max Velocity (m/s)	0.69 (0.10)	0.81 (0.15)	0.059

4. Discussion

Mixed Reality headset holds great promises for enabling portable, self-guided mobility assessment that can be undertaken more regularly without clinician oversight, and subsequently increase the efficiency of current healthcare practice. This investigation is the first to evaluate the validity of MR headset (HoloLens) for mobility assessment in young and older adults. Given the unique advantage of multi-modal user interaction methods(visual/audio/gesture/gaze), this device enables users to initiate and complete a set of valid mobility assessment with step by step guidance, and record head movement as a mean for objective measure of performance.

Overall, our preliminary investigation indicates that it is possible to use a mixed reality headset to automatically guide both young and old user to complete common functional mobility tests (TUG and STS) with good measurement accuracy in comparison to industry standard inertial sensors. More specifically, by comparing the vertical kinematic measures (displacement, velocity, and acceleration) derived from the HoloLens and Opal sensors, we found good to excellent signal agreement for the majority of STS and TUG measures (Xcor 0.74-0.99, NRMSE ~10%), with better signal agreement observed in STS task, as each sit-to-stand cycle allows for displacement calibration adjustment (Zero Displacement Update) [24]. For sensor signal comparison in TUG task, however, due to the integration and drift error associated with IMU sensor and lack of viable calibration during the walking period (10s or longer), the vertical displacement derived from IMU sensors was biased in comparison to HoloLens output (Fig.2e), resulting unsatisfactory NRMSE measures (14.07-19.56%). Moreover, because of HoloLens utilized both depth sensor and IMU sensor to derive the displacement measure (the depth sensor on HoloLens is similar to the Kinect sensor, which has been extensively validated for accuracy in kinematic measure [27, 28]), we would expect the HoloLens output in displacement measure is more trustworthy than IMU derived displacement measure. Although AP displacement comparison between sensors was not analyzed statistically (due to lack of viable drift correction method for IMU sensors), our exploratory investigation found that only HoloLens AP displacement measure matches the standard 3m walking distance utilized in the TUG, whereas the IMU derived AP displacement measure severely underestimate the walking distance (evident by the displacement measure in AP direction as shown in Fig. 3).

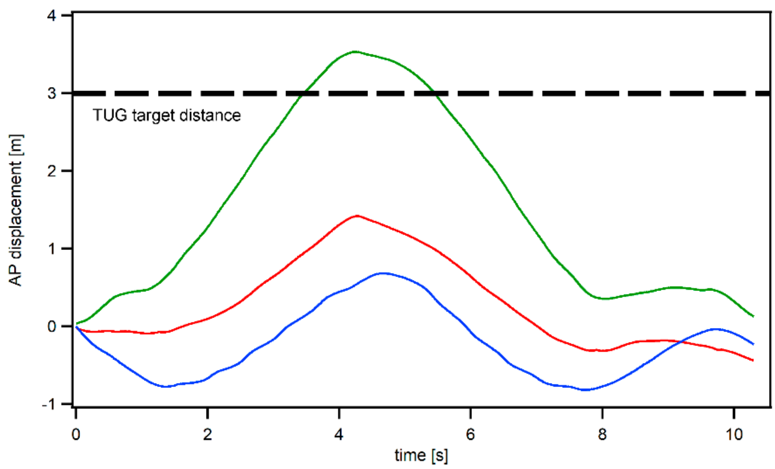

Figure 3. Sample AP displacement in TUG task. Red denotes HD sensor, Blue denotes LB sensor. Note only HoloLens AP displacement measure correctly match the 3m walking path utilized in TUG task.

Additionally, the sensor-derived completion time of STS and TUG was also compared with current gold standard manual stopwatch method using Bland-Altman agreement plot. Excellent agreement with stopwatch timing was found for both STS and TUG completion time, with HoloLens measure demonstrate less than 0.2s in measurement bias and less than 0.8s in range of error).

Because of the relatively healthy nature of the OA participant who completed the test without any assist (only 1 OA participant has significant fall risk, PPA >2), no group difference for sensor-derived performance measure was found between YA and OA, with only marginal difference observed in maximum ascending velocity in STS and TUG task. The two excluded OA participants who have significant fall risk, however, could not complete the STS and TUG task without requiring physical assist, indicating the stand-alone HoloLens device and self-guided functional mobility tests may not be suitable for those who already have severe mobility deficits.

We acknowledge certain limitations for this investigation, most of which related to the pioneering use a novel technology. First, due to the lack of optical motion tracking equipment for portable/community testing, industry standard IMUs were utilized for this validation study. Therefore, due to IMU's inherent integration/drift error and the limited data access to HoloLens displacement data, horizontal (AP/ML) kinematic measurement comparison was not conducted. Secondly, the relatively small sample size, and healthy nature of OA participants who can complete the functional tests without any assist, preclude the investigation to detect the diagnostic power of using head-mount device for fall risk/mobility deficit screening in older adults. Therefore, future studies should incorporate optical motion tracking and larger heterogeneous samples to investigate the use of HoloLens for fall risk screening.

Supplementary Materials: The following are available online, Video S1: Demonstration recording of the MR-based TUG task. Video S2: Tutorial video shown to user on how to put on and adjust the headset.

Author Contributions: conceptualization, R.S and J.S.; methodology, R.S.; software, R.S. and R.A.; formal analysis, R.S.; writing—original draft preparation, R.S.; writing—review and editing, R.S., R.A and J.S.; visualization, R.S.; supervision, J.S.; project administration, J.S.; funding acquisition, R.S. and J.S

Funding: This research was funded by University of Illinois at Urbana Champaign CHART Seed grant

Acknowledgments: The authors would like to thank Microsoft for donating the HoloLens for this research project. The authors would also like to acknowledge Mr. Shrey Pareek for assistance in the software development.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

List of abbreviations

IMU – Inertial Measurement Unit
HMD – Head-mounted Display
MR – Mixed Reality
MoCA- Montreal Cognitive Assessment,
ABC- Activity-specific Balance Scale
STS - Five time Sit to Stand test
TUG – Timed Up and Go test
PPA – Physiological Profile Assessment
RT – Reaction Time
MET – Melbourne Edge Test
Proprio- Proprioception
KneeMax – Maximal isometric knee extension force
AP – Anterior Posterior
ML – Medial Lateral
VT - Vertical
OA - Older Adults
YA – Young Adults
HD – Head IMU sensor
LB – Lower Back IMU sensor
A – Acceleration
V – Velocity
D – Displacement
NRMSE – Normalized Root Mean Squared Error
Xcor – Cross Correlation Coefficient
ZDU – Zero Displacement Update

References

1. Rubenstein, L. Z., Falls in older people: epidemiology, risk factors and strategies for prevention. *Age and ageing* **2006,** 35, (suppl_2), ii37-ii41.
2. Bergen, G., Falls and fall injuries among adults aged≥ 65 years—United States, 2014. *MMWR. Morbidity and mortality weekly report* **2016,** 65.
3. Florence, C. S.; Bergen, G.; Atherly, A.; Burns, E.; Stevens, J.; Drake, C., Medical costs of fatal and nonfatal falls in older adults. *Journal of the American Geriatrics Society* **2018,** 66, (4), 693-698.
4. Ory, M. G.; Schechtman, K. B.; Miller, J. P.; Hadley, E. C.; Fiatarone, M. A.; Province, M. A.; Arfken, C. L.; Morgan, D.; Weiss, S.; Kaplan, M., Frailty and injuries in later life: the FICSIT trials. *Journal of the American Geriatrics Society* **1993,** 41, (3), 283-296.
5. Panel on Prevention of Falls in Older Persons, A. G. S.; Society, B. G., Summary of the updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. *Journal of the American Geriatrics Society* **2011,** 59, (1), 148-157.
6. Stevens, J. A.; Phelan, E. A., Development of STEADI: a fall prevention resource for health care providers. *Health promotion practice* **2013,** 14, (5), 706-714.
7. Buatois, S.; Miljkovic, D.; Manckoundia, P.; Gueguen, R.; Miget, P.; Vançon, G.; Perrin, P.; Benetos, A., Five times sit to stand test is a predictor of recurrent falls in healthy community-living subjects aged 65 and older. *Journal of the American Geriatrics Society* **2008,** 56, (8), 1575-1577.

8. Shumway-Cook, A.; Brauer, S.; Woollacott, M., Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Physical therapy* **2000,** 80, (9), 896-903.
9. Sun, R.; Sosnoff, J. J., Novel sensing technology in fall risk assessment in older adults: a systematic review. *BMC geriatrics* **2018,** 18, (1), 14.
10. Shany, T.; Wang, K.; Liu, Y.; Lovell, N. H.; Redmond, S. J., Are we stumbling in our quest to find the best predictor? Over-optimism in sensor-based models for predicting falls in older adults. *Healthcare technology letters* **2015,** 2, (4), 79-88.
11. Shany, T.; Redmond, S.; Marschollek, M.; Lovell, N., Assessing fall risk using wearable sensors: a practical discussion. *Zeitschrift für Gerontologie und Geriatrie* **2012,** 45, (8), 694-706.
12. Howcroft, J.; Kofman, J.; Lemaire, E. D., Review of fall risk assessment in geriatric populations using inertial sensors. *Journal of neuroengineering and rehabilitation* **2013,** 10, (1), 91.
13. Hsieh, K. L.; Fanning, J. T.; Rogers, W. A.; Wood, T. A.; Sosnoff, J. J., A Fall Risk mHealth App for Older Adults: Development and Usability Study. *JMIR Aging* **2018,** 1, (2), e11569.
14. Hsieh, K. L.; Roach, K. L.; Wajda, D. A.; Sosnoff, J. J., Smartphone technology can measure postural stability and discriminate fall risk in older adults. *Gait & posture* **2019,** 67, 160-165.
15. Kress, B. C.; Cummings, W. J. In *11-1: Invited Paper: Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices*, SID Symposium Digest of Technical Papers, 2017; Wiley Online Library: 2017; pp 127-131.
16. Menz, H. B.; Lord, S. R.; Fitzpatrick, R. C., Acceleration patterns of the head and pelvis when walking on level and irregular surfaces. *Gait & posture* **2003,** 18, (1), 35-46.
17. Psarakis, M.; Greene, D. A.; Cole, M. H.; Lord, S. R.; Hoang, P.; Brodie, M., Wearable technology reveals gait compensations, unstable walking patterns and fatigue in people with multiple sclerosis. *Physiological Measurement* **2018,** 39, (7), 075004.
18. Lord, S. R.; Menz, H. B.; Tiedemann, A., A physiological profile approach to falls risk assessment and prevention. *Physical therapy* **2003,** 83, (3), 237-252.
19. Nasreddine, Z. S.; Phillips, N. A.; Bédirian, V.; Charbonneau, S.; Whitehead, V.; Collin, I.; Cummings, J. L.; Chertkow, H., The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* **2005,** 53, (4), 695-699.
20. Powell, L. E.; Myers, A. M., The activities-specific balance confidence (ABC) scale. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* **1995,** 50, (1), M28-M34.
21. Ejupi, A.; Brodie, M.; Gschwind, Y. J.; Lord, S. R.; Zagler, W. L.; Delbaere, K., Kinect-based five-times-sit-to-stand test for clinical and in-home assessment of fall risk in older people. *Gerontology* **2016,** 62, (1), 118-124.
22. Sun, R.; Aldunate, R. G.; Paramathayalan, V. R.; Ratnam, R.; Jain, S.; Morrow, D. G.; Sosnoff, J. J., Preliminary evaluation of a self-guided fall risk assessment tool for older adults. *Archives of gerontology and geriatrics* **2019**.
23. Sun, R.; Moon, Y.; McGinnis, R. S.; Seagers, K.; Motl, R. W.; Sheth, N.; Wright, J. A.; Ghaffari, R.; Patel, S.; Sosnoff, J. J., Assessment of postural sway in individuals with multiple sclerosis using a novel wearable inertial sensor. *Digital Biomarkers* **2018,** 2, (1), 1-10.
24. Millor, N.; Lecumberri, P.; Gómez, M.; Martínez-Ramírez, A.; Izquierdo, M., Drift-free position estimation for periodic movements using inertial units. *IEEE journal of biomedical and health informatics* **2014,** 18, (4), 1131-1137.

25. Walgaard, S.; Faber, G. S.; van Lummel, R. C.; van Dieën, J. H.; Kingma, I., The validity of assessing temporal events, sub-phases and trunk kinematics of the sit-to-walk movement in older adults using a single inertial sensor. *Journal of Biomechanics* **2016,** 49, (9), 1933-1937.
26. Bland, J. M.; Altman, D., Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet* **1986,** 327, (8476), 307-310.
27. Clark, R. A.; Bower, K. J.; Mentiplay, B. F.; Paterson, K.; Pua, Y.-H., Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. *Journal of biomechanics* **2013,** 46, (15), 2722-2725.
28. Clark, R. A.; Pua, Y.-H.; Fortin, K.; Ritchie, C.; Webster, K. E.; Denehy, L.; Bryant, A. L., Validity of the Microsoft Kinect for assessment of postural control. *Gait & posture* **2012,** 36, (3), 372-377.