Abstract:

Background: Clinical studies in multiple sclerosis (MS) often require accurate measurement of walking distance. Utilisation of electronic devices could theoretically improve this. Mobile devices have the potential to continuously monitor health by collecting movement data. Popular fitness trackers record steps taken and distance travelled, typically using a fixed-stride length. However, applications using fixed-stride length may be less accurate in those with altered gait patterns. While useful for everyday purposes, medical monitoring requires greater accuracy.

Objective: Our aim was to determine the agreement and reliability of using a smart-phone application to measure distance walked.

Method: A phone application (mSteps) was developed and tested in a pilot study and then a validation study, looking at an indoor and outdoor setting with People with MS (PwMS) and a control cohort.

Results: In the pilot study the 95% Limits of Agreement (LOA) for outdoor tracking in control cohort lay within the a priori defined limit, however the indoor tracking in both cohorts did not meet the defined limit. The app was then successfully validated outdoors, in PwMS.

Conclusion: mSteps could be used to accurately measure distance outdoors, in PwMS. There is still a need for solutions to accurately and reliably measure distance walked, indoors.

Background

Measurement of walking distance remains central to clinical trials measurement in multiple sclerosis (MS). The Expanded Disability Status Scale (EDSS) is perhaps the most commonly used disability measure in research trials and clinical practice. It relies on an accurate measurement of distance walked. The gold standard used to measure distance walked is the trundle wheel. However if the patient is unable to perform at full capacity on the day of assessment or there are limitations such as lack of space, walking aids not being present or time constraints, the patient is often asked to provide an estimate of their walking distance. Mobile devices, such as smart phones, have the potential to continuously monitor activity by collecting movement data, including walking speed and distance walked.

There are several mobile devices that can be used to measure walking and provide adjunctive data on patient mobility. The BioStamp® is an example of a novel wireless device which is used to examine gait characteristics of People with MS (PwMS) under controlled situations.[1] Other popular activity monitors include FitBit® bands, ActiGraph and the StepWatch Activity Monitors® (SAM). [2],[3],[4] In one study comparing the Fitbit® Ultra, SAM, Nike Fuelband® and Yamax digi-walker SW701 Pedometer® (devices that monitor stepping activity) the SAM demonstrated greatest accuracy (intra-class coefficient 0.97).[5] It has also been validated as an accurate tool for measuring stride count in PwMS.[6],[7] Another study comparing the Digi-Walker SW-200 pedometer (Yamax), the UP2 and UP Move (Jawbone), and the Flex and One (FitBbit), as well Health app (Apple), Health Mate (Withings) and Moves (ProtoGeo Oy) demonstrated that the waist-worn Fitbit One was the most accurate sensor for measuring steps when walked on a treadmill. [8]

In addition, a number of popular phone applications used for exercise can be used to measure walking/running and these include MapMyRun®, MapMyWalk®, Alpine Quest GPS Hiking®, and the Nike + Running App®.[9]–[11] Whilst the phone applications achieve accurate measurements for walking/running distance they have not been studied in PwMS and there are also issues surrounding the secure and confidential collation of participant data in clinical studies.

Popular fitness trackers record steps taken and distance travelled over time, typically using a fixed-stride length. While useful for everyday purposes, applications using fixed-stride length may be less accurate in those with altered gait patterns, as with PwMS. Medical monitoring requires greater accuracy and validation of the devices in the respective disease being studied. Current generation smartphone hardware contains similar sensors as in medical devices and popular fitness devices and are an obvious target for the development of a measurement device. There are currently no wearable technologies to calculate walking distance and time taken in PwMS, especially a tool that is unaffected by gait disturbances.[12] There is an unmet need for such a tool that would alleviate much of the uncertainty around walking measurements in clinical practice.[13]

Objective

To pilot a smart phone measure in PwMS to determine the agreement and reliability of distance walked using the mSteps mobile phone application, to facilitate EDSS measurement, both indoors and outdoors. The second objective was to validate mSteps as a measure for distance walked in comparison to a gold standard of the trundle wheel.

Method

mSteps development:

We developed the mobile phone application called *mSteps.* The mSteps app has a clock timer and a distance counter, which automatically stops when the person stops moving for more than 20 seconds. Both of these counters restart when the person starts moving again. The application was built on an iPhone platform and an iPhone 6s was used as the study phone. The trundle wheel was used as the gold standard. Our study and application development included significant patient involvement in determining the acceptability and design of the mSteps project.

The app was designed to be tested indoors and outdoors. The outdoor component utilised the phones inbuilt GPS receiver to provide location and time information to the application. GPS does not work well indoors due to the signals from satellites becoming attenuated by roofs, walls and other objects.[14] To address this issue for the indoor component, we used Wi-Fi positioning nodes within an accurately mapped corridor, along with the phones inbuilt 3-axis accelerometer to produce distance walked.

The mSteps application screen shots are shown below in Figure 1:

Figure 1: mSteps application screen shots

The mSteps smart-phone application measured both the time taken to do the walking assessment and the distance walked. For the purpose of the assessment the participant was asked to walk along a corridor, for as long as they comfortably can. The phone application timer uses the smart-phones inbuilt clock.

Study design and participants:

There were two parts: (1) the *pilot* study to determine the reliability and agreement of mSteps and (2) the *validation* study.

The *pilot* study enrolled 25 PwMS, (cohort 1), recruited from studies already being run by the UCL Queen Square MS centre (QSMSC) and 10 people without MS (cohort 2). The *validation* study enrolled 100 PwMS, (cohort 3), again recruited from studies already being run by UCL QSMSC.

In the *pilot* study; the application was trialled indoors on cohort 1 at three different time points. The application was then trialled on cohort 2 indoors and outdoors, referred to as cohort 2(a) and cohort 2(b) respectively, at three different time points. If the *pilot* study results met the *a priori* defined criteria, we planned to proceed to the *validation* study with cohort 3. Study cohorts are visually represented in figure 2.

Figure 2: Study cohorts

Participants that were recruited to cohorts 1 and 3 all had a confirmed diagnosis of MS,[15],[16] were 18 years or older with an EDSS 1.0 - 7.5. Participants with EDSS 7.5 were only recruited if they were able to walk a few steps.[17] Cohort 2 were 18 years or older and had no known neurological or mobility impairment.

Procedure:

- Participant was guided to the walking assessment area and asked to stand/sit still during set-up
- The arm band was positioned on the participant’s arm and the smart-phone application activated
- The participant was asked to walk at a pace comfortable to them without prolonged rest, for 25 feet (*pilot* study) or for as long as they could, at a pace comfortable to them without prolonged rest (*validation* study), whilst a staff member walked alongside them with a trundle wheel

Statistical analysis

Bland and Altman established the use of bias and precision estimates as a standard method when comparing a new method of measurement (here mSteps) against a gold standard (here the trundle wheel).[18]

We therefore used it to determine the agreement between the two methods that measure a continuous variable by calculating 95% limits of agreement (LOA).[18],[19] If the 95% LOA are within the *a priori* clinical acceptable LOA then we can conclude that the two methods of measurement are interchangeable.

In summary, using the method 'where the true value varies – non-constant situation', we model the observed difference between the distance measured by mSteps and trundle wheel using the sum of the: mean difference (bias), a random within subject error (within subject variation) and a random between subjects effect (heterogeneity).[20] The within subject variance of the paired difference between distances measured by mSteps and trundle wheel is calculated using a one-way analysis of variance, using the difference between matched pairs as the dependent variable. The residual mean square is the estimated variance of multiple between-method differences for the same subject.[20] The heterogeneity (between subject variance) is calculated by subtracting the residual mean square from the mean squares for subjects and this is then weighted for the number of participants (n= 35) and the number of observations (n = 3) per subject.[20] Taking the square root of the sum of the within subject variance and heterogeneity provides the standard deviation which can then be used to calculate the upper and lower limits of the 95% LOA: mean difference +/- (1.96 x total variance).

For this approach, the *a priori* clinical acceptable difference is < 1.524m (20% of the prespecified 7.62m). Should the upper and lower bounds of the 95% LOA be less than the prespecified distance of 1.524m, we will assume that the new method (mSteps) is interchangeable with the gold standard trundle wheel.

In the validation study, the distance walked using the app and trundle wheel was captured at a single timepoint, only in MS participants (cohort 3). The a priori defined limit for the validation study, was set as 5m as this was determined to be a reasonable threshold based on the walking distances used to calculated EDSS.[21] If the difference between the app and trundle wheel was less than the *a priori* defined 5 meters, then the app would be seen as a validated tool.

Statistical analysis was completed using R statistical software version 4.0.3.[22]

Results

The pilot and validation study demographics are displayed in Table 1. The trials unit at QSMSC has a particular interest in trials in progressive MS and so necessarily the PwMS had a median EDSS of 6.0 with an age range approximately 50-60 years. To provide information in a younger group, cohort 2 had a median age of around 30 years old.

Table 1: Study demographics

$^{+}$ = Age reported as median (interquartile range)

$^{++}$ = EDSS reported as median (range)

For cohorts 1 and 2(a), mean difference (bias) between the app and trundle wheel was -0.097m and 0.342m respectively (Table 2 and Figure 3). The 95% LOA for cohort 1 (-2.450 to - 2.266) and cohort 2(a) (-2.020 to 2.705) lay outside the *a priori* defined limit of 1.52m (Table 2). For cohort 1 and cohort 2(a), there was a relationship between the magnitude of the measurement and the differences (Figure 3). This relationship could not be removed using log transformation or working with ratios and needs to be taken into account when interpreting the results. The mean difference (bias) for Cohort 2 (b) in the *pilot study* between mSteps and the trundle wheel was -0.013m (Table 2 and Figure 4) and the 95% LOA for this cohort (-0.455 to 0.429) lay within the *a priori* defined limit of 1.52m (Table 2), and therefore we proceeded to capture data with the validation cohort 3.

Table 2: 95% limits of agreement for indoors and outdoors using Bland-Altman repeated measures approach for the pilot study

Figure 3: Bland-Altman plots for cohort 1 and 2(a)

In each plot, the blue demonstrates the mean difference between the app and trundle wheel. The red lines show the 95% limits of agreement.

Figure 4. Bland-Altman plot for cohort 2(b)

The blue line demonstrates the mean difference between the app and trundle wheel. The red lines show the 95% limits of agreement.

Validation study:

The mean difference (bias) between the app and gold standard for cohort 3 (n=100) was 0.262m. The lower LOA was -1.496 (95% CI -1.802 to -1.191) and the upper LOA was 2.020 (95% CI 1.715 to - 2.325). The 95% LOA (and respective 95% confidence intervals) was within the pre-specified acceptable difference of 5m (Table 3 and Figure 5). When stratifying the MS cohort by EDSS (> 6.0 or less), age (> 53.5) and sex, all sub-groups demonstrated acceptable agreement between the app and gold standard (Table 3 & Supplementary Figures 6 a-f).

Table 3: 95% limits of agreement for cohort 3 using Bland-Altman approach

EDSS = Expanded disability status scale, MS = multiple sclerosis

m = metres

Figure 5. Bland-Altman plot for cohort 3

Blue line demonstrates the mean difference (bias) between the app and trundle wheel. The green lines show the 95% limits of agreement. 95% confidence intervals highlighted in dashed red lines.

m = metres

Discussion

The results from the pilot study conducted *outdoors* displayed a very good agreement between the application and the trundle wheel in the control cohort. The MS cohort was not walked outdoors for the pilot study due to worsening weather conditions.

The *indoor* pilot study (cohorts 1 and 2(a)) results showed there was a lack of agreement between the mSteps application and the trundle wheel. This confirms that the indoor walk functionality of mSteps was neither reliable nor accurate. The control cohort also failed to demonstrate agreement between the app and trundle wheel which excludes the possibility that the lack of agreement in PwMS was caused by uneven gait patterns.

The indoor functionality was built using the phones in-built accelerometer and wi-fi positioning. Indoor positioning systems are still being developed throughout the world, and as such there isn't a 'best' solution yet. Most indoor positioning systems are developed for larger spaces, to monitor foot traffic, such as in commercial spaces. Generally, accuracy for Wi-Fi based systems varies between 5 to 15 meters[23] and so comparatively the mSteps indoor has performed on par. There is also a likely relationship between the magnitude of distance walked and the differences between the trundle wheel and mSteps in cohort 1. We postulate this could be for several reasons, such as; the Wi-Fi positioning nodes not being sensitive enough to capture the phones position in relation to the corridor, Wi-Fi signal strength dropping due to frequent signal outages and being in close relation to an MRI machine. However, with the promise of 5G network rollout, we can expect to see solutions using indoor positioning systems improving vastly in the coming years. Further limitations with the solution we used for indoor positioning was the need to map out the floor plan accurately to allow the application to determine where, within the corridor, it was positioned at any one time. Replicating the same, in a solution that is used in a user's home, would be challenging and not feasible. To our knowledge, there are no current devices that are studied and validated for use in an MS cohort, to measure distance walked, indoors. We recognise that there are many mobile devices that can be used accurately, indoors, to measure step count and activity levels.[12]

The mSteps app was then validated in a large MS cohort tested *outdoors* where the 95% LOA between the app and the trundle wheel lay within the pre-specified clinically significant difference of 5m. Furthermore, we demonstrated the validity of the mSteps app in several pre-specified subgroups including EDSS $\geq$ 6.0. An overwhelming number of fitness trackers have shown to be accurate outdoors, by utilising GPS, so our findings were not surprising. Smartwatches and applications such as the Garmin Forerunner[24] and MapMyRun[11] respectively have been widely used to track distances moved outdoors, however none of these devices have been tested and validated for use in PwMS. There are several activity monitors and pedometers that have been studied in MS cohorts, like the FitBit[25] and ActiGraph[26], however these devices measure step count rather than distance walked which is used to calculate disability scores such as EDSS. Limitations to the outdoor solution were few although, we noticed that, when walking under scaffolding, or building works, the results may not be as accurate. There were several subjects that had slightly increased variation in the differences between the trundle wheel and mSteps. Whilst this variation was small, these potential outliers could be explained by GPS satellite signal accuracy being affected by and worsening closer to buildings, bridges and trees. The signal received can also be reflected off buildings or walls causing a decrease in accuracy. [27]

Conclusion

The pilot study showed that the 95% LOA for both the indoor MS and the indoor control cohorts lay outside the a priori defined limit. As both cohorts proved to be not as accurate as the gold standard, the trundle wheel, we can conclude that this is not due to differing gait patterns in the MS cohort. The accuracy of indoor GPS solutions currently available, is not as desirable and therefore we suspect that better indoor GPS modelling may help the accuracy, in the future.

In the pilot outdoors control study, we demonstrated accuracy and reliability of the mSteps phone application and we then went on to validate this in a group of 100 PwMS suggesting that mSteps could be useful in calculating the EDSS outdoors.

Funding

This study was funded by the QSMSC Research Fund, UCLH Charity from donations by patients and their families.

Disclosures

SA has received support from the UK MS Society.

MB Received Support from NIHR Clinical Research Network and the UK MS Society.

CT is currently being funded by a Junior Leader La Caixa Fellowship (fellowship code is LCF/BQ/PI20/11760008), awarded by “la Caixa” Foundation (ID 100010434). She has also received the 2021 Merck’s Award for the Investigation in MS, awarded by Fundación Merck Salud (Spain) and a grant awarded by the Instituto de Salud Carlos III (ISCIII), Ministerio de Ciencia e Innovación de España (PI21/01860). In 2015, she received an ECTRIMS Post-doctoral Research Fellowship and has received funding from the UK MS Society. She is a member of the Editorial Boards of Neurology and Multiple Sclerosis Journal. She has also received honoraria from Roche and

Novartis and is a steering committee member of the O'HAND trial and of the Consensus group on Follow-on DMTs.

In the last 3 years, JC has received support from the Efficacy and Evaluation (EME) Programme, a Medical Research Council (MRC) and National Institute for Health Research (NIHR) partnership and the Health Technology Assessment (HTA) Programme (NIHR), the UK MS Society, the US National MS Society and the Rosetrees Trust. He is supported in part by the NIHR University College London Hospitals (UCLH) Biomedical Research Centre, London, UK. He has been a local principal investigator for a trial in MS funded by the Canadian MS society. A local principal investigator for commercial trials funded by: Actelion, Novartis and Roche; and has taken part in advisory boards/consultancy for Azadyne, Janssen, Merck, NervGen, Novartis and Roche.

Acknowledgments

The mSteps logo was created by Almudena Garcia-Gomez

References:

1. **Moon Y, McGinnis RS, Seagers K, *et al.*** Monitoring gait in multiple sclerosis with novel wearable motion sensors. *PLoS One*. 2017; **12**(2). Available at: https://pubmed.ncbi.nlm.nih.gov/28178288/ [Accessed August 25, 2020].

2. **Sandroff BM**, **Motl RW**. Comparison of ActiGraph activity monitors in persons with multiple sclerosis and controls. *Disabil. Rehabil.* 2013; **35**(9):725–731.

3. **Block VJ, Lizée A, Crabtree-Hartman E, *et al.*** Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. *J. Neurol.* 2017; **264**(2):316–326. Available at: https://pubmed.ncbi.nlm.nih.gov/27896433/ [Accessed August 25, 2020].

4. **Padua L, Pazzaglia C, Pareyson D, *et al.*** Novel outcome measures for Charcot−Marie−Tooth disease: validation and reliability of the 6-min walk test and StepWatch^{TM} Activity Monitor and identification of the walking features related to higher quality of life. *Eur. J. Neurol.* 2016; **23**(8):1343–1350. Available at: https://pubmed.ncbi.nlm.nih.gov/27160471/ [Accessed August 25, 2020].

5. **Fulk GD, Combs SA, Danks KA, *et al.*** Accuracy of 2 activity monitors in detecting steps in people with stroke and traumatic brain injury. *Phys. Ther.* 2014; **94**(2):222–229. Available at: https://pubmed.ncbi.nlm.nih.gov/24052577/ [Accessed August 25, 2020].

6. **Cindy Ng LW**, **Jenkins S**, **Hill K**. Accuracy and responsiveness of the stepwatch activity monitor and ActivPAL in patients with COPD when walking with and without a rollator. *Disabil. Rehabil.* 2012; **34**(15):1317–1322. Available at: https://pubmed.ncbi.nlm.nih.gov/22200080/ [Accessed August 25, 2020].

7. **Schmidt AL**, **Pennypacker ML**, **Thrush AH**, **Leiper CI**, **Craik RL**. Validity of the stepwatch step activity monitor: Preliminary findings for use in persons with parkinson disease and multiple sclerosis. *J. Geriatr. Phys. Ther.* 2011; **34**(1):41–45. Available at: https://pubmed.ncbi.nlm.nih.gov/21937891/ [Accessed August 25, 2020].

8. **Balto JM**, **Kinnett-Hopkins DL**, **Motl RW**. Accuracy and precision of smartphone applications and commercially available motion sensors in multiple sclerosis. *Mult. Scler. J. - Exp. Transl. Clin.* 2016; **2**:2055217316634754.

9. Anon. Presentation - AlpineQuest. Available at: https://www.alpinequest.net/ [Accessed August 25, 2020].

10. Anon. MapMyWalk. Available at: https://www.mapmywalk.com/ [Accessed August 25, 2020].

11. Anon. MapMyRun. Available at: https://www.mapmyrun.com/ [Accessed August 25, 2020].

12. **Alexander S**, **Peryer G**, **Gray E**, **Barkhof F**, **Chataway J**. Wearable technologies to measure clinical outcomes in multiple sclerosis: A scoping review. *Mult. Scler. J.* 2020:135245852094600. Available at: http://journals.sagepub.com/doi/10.1177/1352458520946005 [Accessed August 25, 2020].

13. **Dalla-Costa G, Radaelli M, Maida S, *et al.*** Smart watch, smarter EDSS: Improving disability assessment in multiple sclerosis clinical practice. *J. Neurol. Sci.* 2017; **383**:166–168. Available at: https://pubmed.ncbi.nlm.nih.gov/29246607/ [Accessed August 28, 2020].

14. **. SS**. TERRESTRIAL GPS POSITIONING SYSTEM. *Int. J. Res. Eng. Technol.* 2013; **02**(04):584–589.

15. **Thompson AJ, Banwell BL, Barkhof F, *et al.*** Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* 2018; **17**(2):162–173.

16. **Lublin FD, Reingold SC, Cohen JA, *et al.*** Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology*. 2014; **83**(3):278–286.

17. **Kurtzke JF**. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983; **33**(11):1444–52.

18. **Martin Bland J**, **Altman DG**. STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *Lancet*. 1986; **327**(8476):307–310.

19. **Bland JM**, **Altman DG**. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* 1999; **8**(2):135–160. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10501650 [Accessed April 29, 2019].

20. **Bland JM**, **Altman DG**. Agreement Between Methods of Measurement with Multiple Observations Per Individual. *J. Biopharm. Stat.* 2007; **17**(4):571–582.

21. **Kurtzke JF**. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology*. 1983; **33**(11):1444–1452. Available at: https://n.neurology.org/content/33/11/1444 [Accessed June 10, 2021].

22. **R Core Team**. R: A language and environment for statistical computing. 2017.

23. Anon. Indoor Positioning, Tracking and Indoor Navigation with Wi-Fi. Available at: https://www.infsoft.com/technology/positioning-technologies/wi-fi [Accessed August 28, 2020].

24. Anon. Forerunner 235 | Garmin. Available at: https://buy.garmin.com/en-GB/GB/p/529988 [Accessed August 28, 2020].

25. **Block VJ, Lizée A, Crabtree-Hartman E, *et al.*** Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. *J. Neurol.* 2017; **264**(2):316–326. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27896433 [Accessed August 13, 2017].

26. **Sandroff BM, Motl RW, Pilutti LA, *et al.*** Accuracy of StepWatch and ActiGraph accelerometers for measuring steps taken among persons with multiple sclerosis. *PLoS One*. 2014; **9**(4):e93511.

27. Anon. GPS.gov: GPS Accuracy. Available at: https://www.gps.gov/systems/gps/performance/accuracy/ [Accessed August 28, 2020].