RELIABILITY OF ACCELEROMETER-BASED GAIT ANALYSIS: PROVIDING A BASIS FOR COMPARISON
Monica Martinez\textsuperscript{1,2}, Christopher Wong\textsuperscript{1}, Helena Mentis, PhD\textsuperscript{1}

\textsuperscript{1}Department of Information Technology, University of Maryland-Baltimore County, 1000 Hilltop Circle, 2 Department of Computer Science, New York University, 70 Washington Square South, New York, NY 10012

Abstract
Just as wearable technology is becoming prevalent in society, the possibility for its applications within the realm of healthcare has also experienced significant growth. For our study we will compare data obtained from clinical devices used for gait analysis with non-clinical devices, such as mobile phones and wearable technology, in order to attest to the efficacy of at-home accelerometer-based gait analysis. Our method involves obtaining step counts from five different devices (APDM, Axivity, Actigraph, AndroSensor, and Fitbit) attached at the waist of participants as they walk on a treadmill for four minute intervals at speeds resembling those of patients with impaired gross motor skills (1.3, 1.1, and 0.9ms). Our research aim is to identify the reliability of each device, their limitations and or advantages, and to provide baseline data for accelerometer-based gait analysis testing.

Introduction
Gait analysis testing is a technique used to both assess the severity of diseases that impact gross motor skills, such as Alzheimer's and Parkinson's Disease, and traumas such as Strokes. And yet, clinical gait testing has been underutilized due to a costly, time-consuming process that often yields unclear and unreliable data.[1] However, recent modern wearable fitness devices have achieved a degree of accuracy and reliability which may make them ideal for conducting at-home gait analysis testing as both a low cost and time-saving alternative for patients and clinicians.

Background
Gait analysis testing is used as a method to quantify the mobility status of a certain disorder in order to inform possible treatment options as well as their outcome. Clinical Gait Analysis consists of a patient walking to and from a set distance followed by a clinician parsing the data into Gait Cycles, which are the time between successive foot contacts from the same limb. These cycles are then used to examine Cadence, Stride Length, Step Length, and Step Variability.

Methods
Our goal was to compare the data obtained from the clinical devices used for gait analysis and the non-clinical devices, such as mobile phones and wearable technology (that can utilize built-in accelerometers to measure gait), in order to attest to the efficacy of at-home accelerometer-based gait analysis. The five devices, APDM, Axivity, Androsensor/Mobile Phone (mobile app), Actigraph, and FitBit Ultra, were chosen to to represent the variety of devices available to all types of consumers. (Figure 1) Further the device selection allowed for device categorization that fully encompasses the possible consumer base for these types of devices: clinicians, researchers, and general consumers.
Based on existing literature that demonstrates that a single tri-axial accelerometer can provide relevant data when attached at the waist, each device was fasted to a belt which was then placed on or around the waist area of each participant.[2] These included the Axivity (Khan), Actigraph (Noah), and APDM (Rigsby). Further, in order to also examine non-gait parameters, such as energy expenditure and variability in healthy participants, it was necessary to monitor the walking speeds of each participant through the use of a treadmill. Participants were instructed to walk on a treadmill for 4 minute intervals each with speeds 0.9, 1.1, and 1.3 m/s, respectively in order to simulate the studied gait of Parkinson’s patients.[3][4] Device output was downloaded from each device using their respective proprietary software with the exception of the Fitbit which required the use of a script to download minute by minute data. (Figure 2)

Figure 1: The Devices

Preliminary Results
Although data analysis is currently in its preliminary phase, we can already make some interesting assumptions based on observable data. For starters, the Fitbit device has yielded data that is inconsistent with what can be observed. The accuracy of the step counts produced by the Fitbit, relative to steps observed, is lacking. What we’ve observed was that for the slowest speed (but not exclusively) the step counts were either not counted or significantly lower than observed. Further, when compared to steps that were counted manually based on the video recordings the Fitbit step counts were significantly lower. (Figure 3) One possible reason for the underperformance of the Fitbit is it’s step counting and recognition algorithm. The Fitbit algorithm recognizes steps based on a motion threshold in which it only registers a step if the motion pattern is large enough.[5] Since most of the zero data in our results occurred when participants walked at the slowest speed, 0.9 ms, one can argue that the Fitbit would be the least reliable choice for gait-analysis since the walking patterns of people in need of gait analysis testing are less likely to be recognized by the Fitbit algorithm. One can thus suggest that a more diverse training data set could be more effective for recognizing walking patterns in greater populations, specifically in elderly and physically disabled populations, making the Fitbit device more effective in clinical settings.
Conclusion
Based on the results obtained from the Fitbit device alone we have been able to compare and demonstrate the shortcomings and unreliability of the step counts produced based on the tested speeds, 0.9, 1.1, and 1.3ms. However it is important to note that although the inaccuracy of the step counts produced by the Fitbit were consistently and significantly lower than the expected and observed number, further analysis is still necessary.

Future Work
Following a more thorough analysis of all the device data the goal is to compare the step counts from each device, first against manual step counts and second against each other. We hope to discover the strengths and limitations of each device as well as highlight the possible source of these limitations.

References

Acknowledgments
The work of Monica Martinez is supported in part by the Distributed Research Experiences for Undergraduates (DREU) program, a joint project of the CRA Committee on the Status of Women in Computing Research (CRA-W) and the Coalition to Diversify Computing (CDC), which is funded in part by the NSF Broadening Participation in Computing program (NSF CNS-0540631).