Consumer wristband activity monitors as a simple and inexpensive tool for remote heart failure monitoring

by

Simon Eitan Bromberg

A thesis submitted in conformity with the requirements for the degree of Masters of Health Science in Clinical Engineering
Institute of Biomaterials and Biomedical Engineering
University of Toronto

© Copyright by Simon Eitan Bromberg 2015
Consumer wristband activity monitors as a simple and inexpensive tool for remote monitoring of heart failure patients
Simon Eitan Bromberg
Masters of Health Science in Clinical Engineering
Institute of Biomaterials and Biomedical Engineering
University of Toronto
2015

Abstract
Daily physical activity is a strong indicator of heart failure (HF) severity that could help identify at-risk patients before acute decompensation and prevent late readmissions. However, remote-monitoring systems rarely incorporate any activity data. Consumer devices have made continuous activity tracking inexpensive and unobtrusive, but it is not clear whether such devices are suitable for monitoring patients. In this study 50 stable HF outpatients were given Fitbit Flex wristbands activity trackers to wear for two weeks to validate the use of these devices for remote monitoring. We developed metrics of peak activity based on minute-level step data that correlated with clinical cardiopulmonary exercise tests. These consumer devices were also found to be reliable with seamless data collection. Future studies examining the change of these activity metrics over longer periods time in relation to other indicators of HF will help establish the use of consumer devices as a viable remote monitoring tool.
Acknowledgements

This project would not have been possible, if not for the support and contributions of many individuals, more than can be mentioned here.

First, I offer my gratitude to my supervisor, Dr. Joseph Cafazzo who has supported and guided me throughout my thesis.

I would also like to thank my supervisory committee, Dr. Heather Ross and Dr. Babak Taati, for their guidance and insights.

I am immensely grateful to Mike Walker, the exercise physiologist who conducted all the cardiopulmonary studies and tirelessly answered my questions and offered help throughout the recruitment process.

Thank you to Deanna Socha for her support in recruiting and managing participants during the first half of the collection period.

Thank you to the staff and students of the Centre for Global eHealth Innovation, in particular Melanie Yeung, Alison Bisson, Aarti Mathur, Anthony Soung Yee, Theresa Lee, Vlad Voloshyn, and Stephanie So. Special thanks to Plinio Morita, Shivani Goyal, Kyle Tsang, and Akib Uddin for providing advice and feedback.

I would also like to acknowledge the support and interest from the TELUS Health team, including Suzanne Rochford, Stephane Couture, Zak Bhamani, Dan Sauve, Luc Gagnon, David Legros, and Kathleen Paterson.

This project would not have been possible without the financial support from Healthcare Support through Information Technology Enhancements (hSITE) and Natural Sciences and Engineering Research Council (NSERC).

Thank you to the Institute for Biomaterials and Biomedical Engineering at the University of Toronto, especially Rhonda Marley, Jeffrey Little, and the other people involved in the Clinical Engineering graduate program.

Finally, I would like to thank my friends and family for their support and encouragement throughout my graduate program. Special thanks are due to my amazing wife Brigit Katz for her unwavering patience and support.
# Table of Contents

Acknowledgements........................................................................................................ iii

Table of Contents ................................................................................................................ iv

List of Tables ........................................................................................................................ vi

List of Figures ........................................................................................................................ vii

List of Abbreviations .......................................................................................................... viii

1 Introduction .................................................................................................................... 1
   1.1 Chapter Outline ........................................................................................................... 3

2 Background ...................................................................................................................... 4
   2.1 Heart failure pathology ............................................................................................... 4
   2.2 Heart failure epidemiology and cost ........................................................................... 6
   2.3 Diagnosis and treatment ............................................................................................ 6
      2.3.1 NYHA classification .............................................................................................. 8
      2.3.2 Cardiopulmonary study ......................................................................................... 9
      2.3.3 Timed walking tests ............................................................................................. 10
      2.3.4 Ejection fraction .................................................................................................. 11
      2.3.5 Telemonitoring .................................................................................................... 11
   2.4 Wearable activity trackers ......................................................................................... 12

3 Literature Review ........................................................................................................... 14
   3.1 Telemonitoring in heart failure .................................................................................. 14
   3.2 Activity monitors in healthcare ................................................................................ 15
   3.3 Activity monitors in heart failure ............................................................................. 16

4 Objectives and hypothesis ............................................................................................. 20
   4.1 Gap in knowledge ..................................................................................................... 20
   4.2 Hypothesis ................................................................................................................ 21

5 Activity monitor study .................................................................................................. 22
   5.1 Methodology ............................................................................................................. 22
      5.1.1 Planning and device testing ................................................................................... 22
      5.1.2 Recruitment ......................................................................................................... 24
      5.1.3 Materials ............................................................................................................. 25
      5.1.4 Cardiopulmonary study ....................................................................................... 26
List of Tables

Table 1: Literature review summary ................................................................. 19
Table 2: Study population and exercise data summary ..................................... 29
Table 3: Correlation of predicted values and step metrics ............................... 37
Table 4: Correlations within subgroup where age ≥ 60 ................................... 43
List of Figures

Figure 1: Symptoms of heart failure ................................................................................................... 5
Figure 2: Study kit components ........................................................................................................ 25
Figure 3: Linear correlation of absolute VO$_2$ peak versus watts ..................................................... 30
Figure 4: % predicted relative VO$_2$ peak versus % predicted watts ................................................. 31
Figure 5: Linear correlation of percent of predicted relative VO$_2$ peak versus watts ...................... 31
Figure 6: Age versus relative VO$_2$ peak .......................................................................................... 32
Figure 7: NYHA versus age box plot .................................................................................................. 32
Figure 8: Fractile distribution of single participant’s steps ................................................................. 34
Figure 9: Relative VO$_2$ versus most steps in 10 minutes ................................................................. 36
Figure 10: % predicted watts versus 99$^{th}$ percentile steps in 1 minute ......................................... 37
Figure 11: % predicted watts versus most steps in 10 minutes .......................................................... 38
Figure 12: Exercise test duration versus 99$^{th}$ percentile steps ..................................................... 38
Figure 13: % predicted relative VO$_2$ peak versus average maximum daily steps in 1 minute ....... 39
Figure 14: % predicted watts versus average maximum daily steps in 1 minute .............................. 39
Figure 15: % predicted relative VO$_2$ peak versus average top 10 steps in 1 minute ....................... 40
Figure 16: % predicted relative VO$_2$ peak versus steps in longest walk ........................................ 41
Figure 17: % predicted relative VO$_2$ peak versus % of active minutes at >60 steps/minute ...... 41
Figure 18: % predicted relative VO$_2$ peak versus average daily walking time ............................. 42
Figure 19: % predicted relative VO$_2$ peak versus average daily steps ............................................ 43
Figure 20: Relative VO$_2$ peak versus most steps in 6 minutes ....................................................... 44
Figure 21: % predicted watts versus percentile steps per minute .................................................... 44
Figure 22: Most steps in 10 minutes versus % predicted watts showing NYHA by color ............... 45
Figure 23: NYHA classification versus 99$^{th}$ percentile step minute box and whisker plot ............. 46
Figure 24: ROC curve for classifier based on 99$^{th}$ percentile steps in 1 minute ..................... 46
Figure 25: Box plot of NYHA versus % of active minutes where > 60 steps/minute ....................... 47
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6MWD</td>
<td>6-minute walk distance</td>
</tr>
<tr>
<td>6MWT</td>
<td>6-minute walk test</td>
</tr>
<tr>
<td>AHA</td>
<td>American Heart Association</td>
</tr>
<tr>
<td>API</td>
<td>Application program interface</td>
</tr>
<tr>
<td>AT</td>
<td>Anaerobic threshold</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>CPS</td>
<td>Cardiopulmonary study</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>DBP</td>
<td>Diastolic blood pressure</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>EF</td>
<td>Ejection fraction</td>
</tr>
<tr>
<td>FEV1</td>
<td>Forced expiratory volume in one minute</td>
</tr>
<tr>
<td>FVC</td>
<td>Forced vital capacity</td>
</tr>
<tr>
<td>HF</td>
<td>Heart failure</td>
</tr>
<tr>
<td>HR</td>
<td>Heart rate</td>
</tr>
<tr>
<td>IBBME</td>
<td>Institute for Biomaterials and Biomedical Engineering</td>
</tr>
<tr>
<td>ICD</td>
<td>Implantable cardioverter-defibrillator</td>
</tr>
<tr>
<td>LVEF</td>
<td>Left ventricular ejection fraction</td>
</tr>
<tr>
<td>MEMS</td>
<td>Microelectromechanical systems</td>
</tr>
<tr>
<td>MET</td>
<td>Metabolic equivalent</td>
</tr>
<tr>
<td>NYHA</td>
<td>New York Heart Association</td>
</tr>
<tr>
<td>O2 sat</td>
<td>Blood oxygen saturation</td>
</tr>
<tr>
<td>REB</td>
<td>Research Ethics Board</td>
</tr>
<tr>
<td>RER</td>
<td>Respiratory exchange ratio</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operating characteristic</td>
</tr>
<tr>
<td>SBP</td>
<td>Systolic blood pressure</td>
</tr>
<tr>
<td>TGH</td>
<td>Toronto General Hospital</td>
</tr>
<tr>
<td>TWT</td>
<td>Total walking time</td>
</tr>
<tr>
<td>UHN</td>
<td>University Hospital Network</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>VCO₂</td>
<td>Rate of carbon dioxide production</td>
</tr>
<tr>
<td>VE</td>
<td>Minute ventilation</td>
</tr>
<tr>
<td>VO₂</td>
<td>Rate of oxygen consumption</td>
</tr>
</tbody>
</table>
1 Introduction

Heart failure (HF), where the heart is unable to meet the demands of the tissues\(^1\), has been described as a growing global epidemic.\(^2\)-\(^4\) HF affects over 23 million people worldwide and costs healthcare systems billions of dollars, primarily from hospitalizations.\(^3\) HF is a long-term chronic condition that significantly restricts a patient’s ability to perform activities of daily living. Prevalence and incidence of HF is rising, mainly due to an aging population and improved survival due to better treatments. Stable patients may experience acute decompensation, where symptoms of HF such as fatigue, shortness of breath, and swelling, worsen and require unplanned office visits or readmission to a hospital.\(^8\) HF has the highest readmission rates among hospitalizing conditions, and repeated hospitalization is a strong predictor of mortality.\(^5,6\)

Decompensation and subsequent hospitalization can be influenced by non-adherence to medication and recommended diet, inability to recognize signs and symptoms of HF, or limited access health care professionals.\(^7\) Signs and symptoms of decompensation can present as much as a week before admission; so earlier intervention may help reduce readmissions.\(^8\) In-person disease management programs have shown promise for reducing readmissions, but these can be expensive and impractical for patients who live far from their clinic.\(^7\) Instead, patients can be taught to engage in self-monitoring, and telemonitoring systems can be used to track patient status using physiological data collected on a daily basis and provide support when needed.\(^7\)

Modern telemonitoring systems can incorporate tools for patients to report symptoms and automatically track key indicators like blood pressure and weight. Research has shown that these systems are effective at improving patient quality of life and reducing costs, readmissions, and mortality.\(^7\) They help patients get attention from clinicians when they need it most, even if they live far away from their clinic. However, these systems currently lack the ability to track a critical aspect of monitoring HF: physical activity.

Limitation of physical activity, due to reduced cardiac output and thus less oxygenated blood reaching skeletal muscles, is one of the first symptoms of HF and often the principal reason for seeking care.\(^1\) Exercise testing, a routine part of a visit to any cardiac clinic, assesses patients’
peak and functional exercise capacity. However, this brief and infrequent snapshot of a patient’s state is rarely reflective of their physiological state at home.9 Furthermore, patients also consistently overestimate their physical activity.10 Research has shown that daily activity monitoring is not only a simpler and a better representation of patient life, but also that it may outperform clinical exercise tests for predicting mortality and hospitalization.33 Although research into remote activity monitoring in HF has been ongoing for decades and there are now many different activity monitors on the market, activity monitoring has yet to transfer from research to clinical practice and telemedicine.11,48 This may be because of the impractical nature of the devices used in these studies, which create obstacles for adoption in practical applications. For example, many papers use devices that are worn on the belt or upper arm, which are obtrusive because they are bulky or need to be removed when changing clothing. They are also usually more difficult to acquire than simply ordering online or purchasing off the shelf of an electronics store. These devices also may require technical training and significant time and expenditures if they were to be integrated into existing clinical workflows. Furthermore, the aforementioned studies required their participants to wear the devices for a certain number of hours per day, or to perform specific activities. While this may be tolerated for short-term studies, an activity monitoring approach that analyzed a patient’s normal activity—rather than requiring them to adjust their habits—would be preferred.

In this project, I seek to validate the use of a consumer wristband activity monitor for the remote monitoring of physical activity in HF patients. Modern consumer devices are easy to acquire, inexpensive, and simple to use. They collect data wirelessly using a mobile phone and can provide step counts not only as daily totals, but down to the minute level as well. This thesis will show that physical activity as measured by a wristband activity tracker correlates with HF severity via standard cardiopulmonary exercise testing. The results will be used to optimize simple metrics based off of minute-by-minute step data for this correlation that could be used in practical applications. We will also show that this approach has significant advantages that will aid with practical adoption, including that it will not require perfect patient adherence or the imposition of daily walking requirements to be useful. The study will also provide evidence for reliability and usability of this consumer technology as a telemonitoring tool by simulating how these tools would be used in a remote monitoring application.
1.1 Chapter Outline

In Chapter 2 of this thesis, I present background information about HF and the concepts related to this project. This includes a discussion of the definition of HF, its symptoms, diagnosis, and management as well as its impact and cost. There is also a discussion of wearable technology and how wearable devices relate to healthcare. Chapter 3 provides a review of literature on telemonitoring and the use of wearable activity monitors in healthcare. There is a section specifically focused on studies of HF and daily activity monitoring. Chapter 4 summarizes the gaps in knowledge that provide the justification for this thesis project, as well as the objectives and hypothesis of the present work. Chapter 5 presents the methodology and results of the HF activity monitoring study conducted as part of this thesis project. Chapter 6 contains a discussion of the findings within the context of previous work and this study’s limitations. Chapter 7 concludes this document with a summary of contributions, and discusses implications of the present research as well as directions for future work.
2 Background

2.1 Heart failure pathology

Heart failure (HF) is a chronic condition in which the pumping of the heart cannot keep up with the changing needs of the body. To meet increased oxygen demands during physical exertion, the heart pumps faster and harder. When the heart is impaired, it has to work even harder to maintain circulation, until it reaches a point where it is unable to keep up and tissues will begin to feel fatigue. HF often starts on the left side of the heart, where the left-ventricle either cannot contract to push blood out normally (systolic failure), or it cannot relax and fill normally (diastolic failure) due to muscle stiffness. Left-sided HF causes pulmonary edema and shortness of breath.

Right-sided HF often occurs as a result of left-sided HF, when the right ventricle, which pumps de-oxygenated blood returning through the veins to the lungs, is damaged. Congestive HF occurs when the reduced blood flow out of the heart causes blood vessels and tissues throughout the body to become engorged due to fluid backup (Figure 1). This causes shortness of breath and edema, especially in the legs, ankles, liver, and abdomen. If not managed properly, HF patients can experience a sudden or gradual worsening of their symptoms that results in acute decompensation and may require hospitalization or a visit to their physician. Rapid weight gain due to fluid retention is a very common sign of decompensation.

Patients usually experience limitations on their physical activity, which varies based on the severity of their condition. At rest, the heart may be able to meet circulation demands and a patient may not experience any symptoms, but as soon as they experience even a slight exercise load, they can become fatigued quickly and feel pain. In severe cases, an individual may experience discomfort even while at rest. Exercise intolerance, as defined by the American Heart Association (AHA), is “the reduced ability to perform activities that involve dynamic movement of large skeletal muscles because of symptoms of dyspnea or fatigue.” It occurs due to inadequate blood flow to skeletal muscles because of impaired cardiac output. Daily activities like walking, climbing stairs, carrying, and physical exertion may result in quick fatigue, pain, shortness of breath, or even acute events.
Heart failure is the end stage of many cardiac diseases, which impair the heart through various forms of damage. For example, coronary artery disease, one of the most common cardiac diseases, is characterized by narrowing of arteries due to atherosclerosis. This causes reduced oxygenation and weakening of heart tissue, as well as myocardial infarction if plaques rupture and block blood flow. Other sources of damage to the heart include hypertension, valvular disease, atrial fibrillation, heart valve disease, congenital defects, and cardiomyopathy from substance abuse. Initially the heart compensates for its weakened state by growing more muscle, stretching, increasing heart rate, narrowing vessels to maintain blood pressure, and diverting blood from less critical tissues. These adaptations mask the symptoms of HF, so a patient may not seek treatment until the disease progresses to the point where it significantly affects daily living.

Figure 1: Symptoms of heart failure
2.2 Heart failure epidemiology and cost

Worldwide, over 23 million people have HF, including over 500,000 Canadians and 5.8 million Americans.\textsuperscript{3,17} HF is strongly prevalent in the elderly, affecting 6–10\% of people over the age of 65 years.\textsuperscript{8} HF tends to affect men more than women.\textsuperscript{3}

There are 825,000 new cases in the USA each year, and the lifetime risk of developing HF is 20\%.\textsuperscript{3,18} There are 1 millions hospitalizations each year in the United States with a primary diagnosis of HF, and it is the most common diagnosis for hospital admissions in patients over 65.\textsuperscript{8} The incidence of HF has continued to increase due to an aging population and prolonged survival thanks to improving treatments.\textsuperscript{8,7}

Although mortality rates have improved over the past few decades, mortality is still as high as 30\% after 3 years of initial HF diagnosis.\textsuperscript{3,19} Readmission rates in HF are the highest among hospitalizing conditions, with 27\% of patients readmitted within 30 days, and 50\% within 6 months.\textsuperscript{9,20} Only 15\% to 20\% of HF hospitalizations are for new diagnoses.\textsuperscript{8} High readmissions are often attributed to challenges faced by patients, namely limited access to specialized healthcare professionals, limited hospital resources, difficulty recognizing signs of decompensation, and failure to adhere to recommended medication, diet, and exercise.\textsuperscript{7} The number of hospitalizations has also been shown to be a strong predictor of mortality.\textsuperscript{6} Furthermore, up to two-thirds of hospitalizations may be preventable.\textsuperscript{21} Early symptoms of decompensation are present as much as 1 week prior to acute decompensation events, so earlier intervention may help reduce hospitalizations.\textsuperscript{8}

In the USA, the total cost of HF was estimated in 2012 to be $30.7 billion, with 80\% related to hospitalization.\textsuperscript{3,8,18,22} HF represents 2\% of the total healthcare expenditures of Western countries.\textsuperscript{19} The total cost of HF is projected to rise to $70 billion by 2030.\textsuperscript{22}

2.3 Diagnosis and treatment

Identifying HF can be difficult because symptoms may be masked initially by the heart’s adaptations, and many features of HF are not organ-specific.\textsuperscript{3} Patient history, physical examination, blood tests, exercise tests, and imaging are all used to identify and monitor HF.

To evaluate HF severity, one must test the heart not only at rest, but also during physical activity.
The heart may be able to meet demands of bodily tissues at rest, but it may fail to keep up under even a slight exercise load. In fact, resting heart measurements, such as ejection fraction, are only weakly correlated to exercise tolerance.\(^1\) Many of the tests done in a clinic are focused on assessing maximal physical activity and exercise capacity. Exercise training also has important beneficial effects on patient outcomes, and it is important for patients to try to stay physically active.\(^2\) One study found that by simply increasing walking duration from 10 to 60 minutes per day, HF patients could improve exercise capacity and their general well being.\(^3\) Numerous studies have shown that exercise-training programs improve exercise capacity in terms of both duration and peak rate of oxygen consumption (\(\text{VO}_2\text{max}\)). Many clinics offer teaching and support programs to help patients stay active, but it is important for patients to engage in self-care and avoid a sedentary lifestyle. However, feelings of hopelessness can overwhelm patients,\(^4\) and they typically have comorbidities that interfere further with physical activity.\(^5\)

There is no cure for HF, but it can be managed effectively and stabilized with diet and lifestyle changes in conjunction with medical treatment.\(^6\) Treatment depends heavily on patient etiology as well as severity and progression of the illness. Many patients take medication to control blood pressure, such as angiotensin converting enzyme inhibitors, diuretics, and beta-blockers.\(^7\) Patients are instructed to adjust their lifestyle and diet to limit further damage to the heart and give it an opportunity to heal, to whatever extent is possible. These adjustments include reducing smoking, alcohol and drug consumption, eating healthier, controlling fluids and salt, and exercising regularly.\(^8\) Many patients receive implants such as a pacemaker, an implantable cardioverter-defibrillator (ICD), or a ventricular assist device.\(^9\) Patients may also undergo operations such as valve replacements, bypasses, angioplasty, or heart transplant.\(^10\)

Stable patients will come in for regular check-ups with their cardiologist, but usually no more than a few times per year. More frequent visits are not feasible for many patients, as the closest cardiac clinic may still be several hours away. Patients may visit their doctor sooner if they experience worsening symptoms, or they may be hospitalized in the event of acute decompensation. During their appointments they undergo blood, imaging, ECG, and exercise tests in addition to physical examinations and discussions with their doctor. They also receive any updates to prescriptions for their HF medication.
2.3.1 NYHA classification

The New York Heart Association (NYHA) Functional Classification is a simple system for classifying self-reported severity of heart failure based on limitations to physical activity. Patients are assigned to one of four classes ranging from I to IV (with I being the least severe and IV being the most severe). The description of the classes below are taken from the AHA website.  

I. No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, dyspnea (shortness of breath).

II. Slight limitation of physical activity. Comfortable at rest. Ordinary physical activity results in fatigue, palpitation, dyspnea (shortness of breath).

III. Marked limitation of physical activity. Comfortable at rest. Less than ordinary activity causes fatigue, palpitation, or dyspnea.

IV. Unable to carry on any physical activity without discomfort. Symptoms of heart failure at rest. If any physical activity is undertaken, discomfort increases.

The classification assigned to a patient is typically based on their clinical interview during a visit. A clinician may ask questions about how the patient feels when he or she gets up to go to the bathroom, climbing stairs, or walking outside. Together with more quantitative test results, a cardiologist can determine how the patient is faring and this guides decisions about treatment and management. The doctor can also quickly compare classifications assigned on previous visits to monitor patient stability and response to treatment.

While NYHA classification is a useful summary of a patient’s functional status and is based on trained medical opinion, it still relies on subjective self-reporting by the patient and experience-based interpretation of words like “mild” versus “marked” by the clinician. Experienced clinicians may consider patient test results when assigning a classification, in case the patient is understating their symptoms. Patients consistently overestimate their physical activity and may have difficulty describing their physical limitations if they are sedentary for most of the day. At the TGH clinic, physicians often use ambiguous classifications, such as “I/II”, or “II/III” to classify patients that do not clearly fit into one classification. In some cases physicians did not even assign a class, or they do not explicitly record it in the patient’s chart. Interpretation of the classes and standards for their use may be different at other clinics, but this merely emphasizes
the limitations of this tool as a classifier. Fortunately, more objective tools for evaluating HF severity are available to supplement this interview-based assessment method.

2.3.2 Cardiopulmonary study
A cardiopulmonary study (CPS) or cardiopulmonary exercise test measures the response of the body to exercise.\textsuperscript{29} Measurements taken at rest cannot predict exercise capacity.\textsuperscript{29} This test involves incremental exercise loads—such as on a stationary bike with increasing resistance—and is symptom-limited, meaning the test is stopped when the patient is fatigued or experiencing pain or discomfort. During the test the patient wears electrocardiogram (ECG) leads, pulse oximetry, a mask over the mouth to rapidly collect and analyze gases inhaled and exhaled during breathing, and a blood pressure cuff so blood pressure can be measured at regular intervals. The test compares how a patient performs relative to their rest state in terms of blood pressure, heart rate, blood oxygen saturation, and respiration. The test also compares the exercise data relative to standardized results according to weight, sex, age, and height.\textsuperscript{30} Thus a patient’s exercise capacity can also be expressed as a percentage of their predicted level.

The key output of the CPS is VO\textsubscript{2} max, or maximal oxygen consumption. VO\textsubscript{2} max is defined as the point at which rate of oxygen consumption plateaus during incremental exercise. Past this point, oxygen consumption does not increase with exercise intensity. Because of physical limitations experienced by HF patients, the actual maximum VO\textsubscript{2} achieved during a CPS is often not a true VO\textsubscript{2} max and is instead referred to as VO\textsubscript{2} peak.\textsuperscript{29} VO\textsubscript{2} max is widely accepted as the best single measure of cardiovascular fitness and maximal aerobic power, and thus is the best indicator of a patient’s exercise capacity.\textsuperscript{31} It can be measured as an absolute value, such as in litres (of oxygen) per minute, but for comparison between individuals it is usually expressed relative to weight, in millilitres per kilogram per minute. A larger person will consume more oxygen than a smaller person. VO\textsubscript{2} max also tends to decrease with age, and women tend to have lower values than men of the same age.

Exercise capacity is reduced even in mild HF.\textsuperscript{29} Reduced exercise capacity is strongly associated with increased heart failure severity, namely higher NYHA classification, worse symptoms, poor quality of life, and decreased survival.\textsuperscript{29} Exercise capacity is assessed regularly during patient visits and reveals a lot about how the patient is feeling.
A sample of the CPS summary reports used in the clinic at TGH can be found in Appendix A. It includes measurements taken at rest, before the test, as well as peak activity and respiratory data. Systolic and diastolic blood pressure (SBP, DBP), heart rate (HR), and blood oxygen saturation (O2 sat) are measured at rest and continuously (or periodically for blood pressure) until the peak is reached and the test is terminated. The drop of HR in 1 minute after the peak is also reported. FEV1 is the forced expiratory volume in 1 second, and FVC is the forced vital capacity. Percent of predicted values are also reported for these indicators of pulmonary function. Key exercise values are reported along with predicted results, which provide normalization to correct for factors like age and sex. The test includes several other measures of cardiopulmonary performance. Peak VE is the peak minute ventilation, the volume of gas exhaled per minute. Peak VCO₂ is the peak carbon dioxide production and the ratio of VE and VCO₂ are reported at anaerobic threshold (AT) and at the peak. AT is the point at which lactic acid starts to accumulate in the blood, and is reported as a raw value as well as relative to measured and predicted peak VO₂. RER is the respiratory exchange ratio, which is the ratio between the amount of oxygen consumed and carbon dioxide produced in one breath. Each of these pieces provide useful information about the patient’s status, but the VO₂ max is the key overall indicator of exercise capacity, and the percent of predicted values provide a quick indicator of the function relative to normal values.

2.3.3 Timed walking tests
A timed walking test involves walking continuously on a set path for a set duration while maintaining a comfortable but near maximal pace. The 6-minute walk test (6MWT) is the most popular. Patients walk back and forth along a 100-ft flat, hard surface for 6 minutes, usually back and forth in a corridor.³² This short test has been shown to perform as well as longer walking tests, and is easier to administer and more comfortable for patients.³² It does not require hooking up the patient to multiple expensive apparatuses like the CPS or supervision by an exercise physiologist. 6MWTs represent a simple measurement of functional capacity and a summary of the response of the body to exercise.³² Unlike the CPS, the workload in a self-paced 6MWT is chosen by the patient, and thus may be more indicative of symptomatic impairment.³³ Although symptom-limited exercise tests using laboratory equipment allow close control and generate plenty of data, they do not represent a typical form of exercise.³³ In the 6MWT patients choose their own workload, which may be more representative of a patient’s functional status.³³
However, despite their simplicity these tests still require administration and supervision by trained personnel and should be performed on a pre-set path with a level surface that is free of obstacles.\textsuperscript{34}

Research has shown that the 6MWT is a satisfactory measure of functional capacity and activities of daily living in patients with heart disease.\textsuperscript{35} It is an objective measure of submaximal exercise that does not rely on patient self-reporting, yet it only requires a walking surface and a timer. The primary variable of interest in the 6MWT is the total distance walked (6MWD). Pulse oximetry may also be measured during the test. Healthy subjects’ 6MWD range from 400–700 metres, although a low 6MWD is nonspecific and nondiagnostic.\textsuperscript{35} Nevertheless, the 6MWD independently predicts HF severity, hospitalization, and mortality.\textsuperscript{34}

2.3.4 Ejection fraction

Ejection fraction (EF) is the percentage of blood by volume that exits the heart when it contracts, usually measured in the left ventricle because it pumps blood to the rest of the body. EF is typically measured with echocardiography, but it can also be measured using other imaging techniques, such as a computed tomography (CT) scan.\textsuperscript{3,36} EF is an important measurement of heart function and HF diagnosis and monitoring.\textsuperscript{36} An EF between 55% and 70% is considered normal.\textsuperscript{36} Below 35% is reduced and generally considered to be HF. Over 50% of HF patients have preserved EF (also referred to as diastolic HF), where the heart contracts normally but cannot fill properly because the tissue is stiff.\textsuperscript{3,36} The distinction between the two is important because they are managed in different ways.\textsuperscript{3}

2.3.5 Telemonitoring

HF patients typically only visit their cardiologist a few times per year. Understanding how a patient is progressing depends largely on intermittent snapshots from these infrequent visits, where patients report on their symptoms and perform various tests. Clinics use telemonitoring systems to help track patient status between visits, and to help identify problems before they become emergencies. Even regularly scheduled phone calls to patients by nurses have been shown to reduce hospitalizations.\textsuperscript{26} Telemonitoring technologies can take many forms, including video-consultations, mobile, automated-device based, interactive voice response, and web-based.\textsuperscript{37} Some systems require patients to manually enter vital signs through a web-portal,
mobile application, or over the phone. Others use modern devices that can automatically transmit data from devices that measure physiological metrics such as weight, heart rate, ECG, and blood pressure.

2.4 Wearable activity trackers

Research involving daily physical activity originally used mechanical pedometers, which have moving parts to track steps, such as a ball that hits a counter switch each time it moves back and forth. These devices typically are worn on the hip and the recorded step count often must be written down and reset each day. Within the past decade, new advancements have enabled the recent surge of new wearable devices that can track health-related metrics. Wearable devices have grown incredibly popular in the consumer market: over 32 million wearable activity and health devices were sold in the US in 2014, and the number continues to rise each year. The rise in wearable technology has popularized the concept of the quantified self, or the idea that a combination of wearable sensors can quantify metrics related to well being. Consumer wearable devices can track a wide variety of metrics relevant to health, including stair climbing (using sensors that measure changes in air pressure), heart rate (using light-emitting diodes and photodetectors), and even skin temperature.

Modern activity trackers use microelectromechanical systems (MEMs) inertial sensors to track movement. The displacement of a tiny mass is converted to an electric signal using piezoelectric components. Thus, modern activity trackers follow the same principles as their pedometer ancestors, but on a much smaller scale, and with significantly more detail. Accelerometer chips usually have multiple sensors arranged to measure acceleration along multiple perpendicular axes. Activity trackers track not only daily totals, but also activity down to the minute level, or even raw accelerometer data with a resolution even lower than one second. Flash memory provides ample data storage that fits in the small space available, is durable, and does not require a lot of power. Other technologies that have contributed to today’s wearable revolution include smartphones, low energy Bluetooth wireless communication, improved batteries, and new and more accessible tools for hardware prototyping and manufacturing. Most of these devices communicate wirelessly with smartphones or computers to upload, analyze, and visualize the data. They come with a rechargeable battery with a life that varies significantly depending on the device’s capabilities and intended use. The Apple Watch cannot go a full day without being
recharged, while Jawbone’s UP MOVE has a replaceable coin-cell battery that lasts up to 6 months. Devices without wireless synchronization have even longer battery lives. If a device needs to be recharged too frequently it becomes very inconvenient for the wearer. This is especially true if the device’s battery does not even last a full day. Typically, batteries inside wristband activity trackers can last several days, and they can be recharged using a special cable that comes with the device.

There are many different accelerometer-based activity trackers on the consumer market, as well as some specifically designed for research applications. There are activity trackers that are worn on the foot, ankle, waist, upper arm, wrist, torso, around the neck, on the finger, and even on the ear. Although placement on the ankle or waist helps to avoid an erroneous counting of steps due to other body movements, a wristband is advantageous because people are already accustomed to wearing items around their wrists, and it does not need to be removed with clothing. Validation studies have compared wristband activity trackers (including the Fitbit Flex) to ones worn on the waist or arm, and shown that these trackers are quite accurate, even compared to research-quality devices. For this project, we specifically focused on consumer-level wristband devices. Some of the devices considered included the Jawbone UP, the Withings Pulse, the Misfit Shine, the Basis Band, and the Fitbit Flex. Each of these devices costs $100 to $200, have slightly different designs and features, and transmit data wirelessly to a smartphone application.


3 Literature Review

3.1 Telemonitoring in heart failure

Primary research, reviews, and reviews of reviews have explored telemonitoring in HF at great length, evaluating different tools and the effect they have on outcomes, cost, and quality of life.\(^{37}\) While evidence showing the benefits of phone-based patient support from medical staff is clear, research on remote monitoring is more controversial.\(^26\) Two large studies of telemonitoring in HF were both unable to demonstrate a significant impact of telemonitoring on rehospitalizations or mortality.\(^5,26\) The failure of these studies, however, could be attributed to the complexity of home monitoring interventions, since individual limitations of the systems employed in those studies may have been too difficult or inconvenient for patients and clinical staff, due to tedious input systems or adding unnecessary layers of communication and complexity.\(^5\) Telemonitoring involves many different coordinated layers. Data must be transmitted at appropriate times, received by clinical staff for analysis so they can send directions back to the patient. The patient must then be able to act on the instructions they receive, which requires proper training from clinical staff.\(^5\) Given the wide variety of tools and methodologies possible for a telemonitoring study, and the many extraneous variables involved, it is not surprising that some research has been inconclusive about telemonitoring. Telemonitoring tools and systems must be developed with usability for both clinical and patient users in mind.

In 2012 Seto et al performed a randomized control trial of a mobile phone based telemonitoring system for HF.\(^39\) 50 patients took daily weight and blood pressure readings and weekly ECGs, and answered daily questions about symptoms for 6 months. Cardiologists received alerts through the system, which often resulted in the patient being contacted. Patients had high adherence rates and found the system improved quality of life through improved self-care and clinical management.

Maeng et al studied the impact of a new telemonitoring system on over 500 members of a health plan over a period of about 5 years.\(^40\) The telemonitoring system included a Bluetooth weight scale with an interactive voice system to collect answers responding to questions about symptoms. The system reduced contributed to significant reductions in hospitalizations, with 11% cost savings and a substantial return on investment.
A 2015 systematic overview of 15 reviews by Kitsiou et al found that home telemonitoring interventions reduced mortality and hospitalization, especially in device-based and mobile telemonitoring. A 2014 review of 32 studies by Grustam et al found that although some studies showed telehealth interventions are cost saving and slightly more effective than usual care, the quality of these studies was low and relevant cost data was lacking in many cases.

While telemonitoring interventions incorporate modern devices for wirelessly transmitting physiological data including weight and blood pressure, few, if any, of the telemonitoring papers discussed the use of activity monitoring tools.

### 3.2 Activity monitors in healthcare

Recent studies have validated the accuracy and reliability of consumer activity monitors, including the Fitbit Flex. Bai et al compared the performance of Fitbit Flex and four other popular devices against two research monitors and a portable metabolic system monitor. Their study, which included 52 participants who each performed varying activities while wearing the devices, concluded that the Fitbit has comparable accuracy for energy expenditure estimation to the research monitors.

Diaz et al validated the use of consumer activity trackers by studying 23 participants while they performed a treadmill exercise with different walking speeds while wearing three hip-based Fitbit One devices and two wrist-based Fitbit Flex devices. They found that the hip devices outperformed the wristband Fitbit Flex in terms of accuracy against observed step counts. However, the accuracy of the Fitbit Flex was still quite high, and was even on par with hip devices for estimated energy expenditure. Note that the present study did not require a perfectly accurate measure of each participant’s step count, but rather a consistent estimate of physical activity.

Several other studies have shown that commercial activity monitors can be used for a variety of applications, such as chronic obstructive pulmonary disease, and transfemoral amputations. A 2013 Mayo Clinic study by Cook et al showed Fitbit trackers worn on the ankle could predict length of stay in older patients following major surgery. Computers inside the patient rooms were interfaced with Fitbit Bluetooth adapters to collect data wirelessly throughout the study.
They found a significant relationship between the number of steps taken during the early recovery period, length of stay, and status at discharge. They also reported that the wireless monitoring was easy and practical. This work provides strong support for the potential of using consumer devices for clinical applications.

### 3.3 Activity monitors in heart failure

Studies correlating daily physical activity measured by activity trackers have been conducted since at least the early 1990s. While a CPS and a supervised 6MWT are different approaches to assessing a patient’s exercise capacity, their normal daily activity is probably even more indicative of impairment. These studies vary in terms of the technologies used and the length of time studied, but they all agree on the value of monitoring HF patients’ daily physical activity by relating it to supervised clinical exercise tests, NYHA classification, or patient outcomes. This literature is discussed below and is also summarized in Table 1.

Davies et al (1992) studied 20 young healthy controls, as well as 30 stable HF patients against an additional 20 age-matched normal controls. HF patients wore the movement monitors, which used mercury switches, on their wrist and the ipsilateral ankle, and had to read the output at the same time everyday for up to 14 days. The movement monitors correlated with questionnaire-based measures of physical activity as well as exercise tests. They predicted that monitors of movement would continue to prove useful in research.

Oka et al (1993) studied 45 patients in treadmill exercise tests followed by two days of normal activity. They used a device worn on the belt with sensors to measure heart rate as well as arm and leg motion. The combination of heart rate and motion data enabled them to distinguish physical activity from small movements of the limbs. Subjects kept a log of their activities and symptoms. They concluded that daily activity in HF patients is low, and that there is a gap between exercise capacity (as measured by clinical tests) and daily physical activity.

Walsh et al in the United Kingdom studied 84 patients between 1987 and 1993 from an initial assessment until death, transplantation, or the end of the study period (on average approximately two years). Upon study entry, participants performed treadmill exercise tests, a corridor-walking test, and a week of daily activity tracking using two pedometers worn on the hips. While
neither an incremental nor a fixed workload exercise test predicted prognosis, reduced weekly pedometer scores did. They conclude that daily activity levels are a strong predictor of death and may outperform laboratory-based exercise testing.

Jehn et al (2009) assessed 50 HF patients with a CPS and a 6MWT, followed by 7 days of monitoring using an accelerometer sensor worn on the belt for at least 12 hours per day. Data was copied from the devices upon return. They found total daily walking time (TWT) correlated with VO\textsubscript{2} peak and 6MWD, (r \approx 0.7). TWT and time spent in fast walking mode proved to be the most successful at correctly classifying patients in NYHA class III. Fast walking was defined as speeds from 83 to 115 m/minute, and was analyzed relative to the daily total wearing time. A 2013 paper from the same group used accelerometer sensors capable of transmitting data remotely via mobile network. 155 patients wore a device on the hip attached to the belt and was about the size of a matchbox. They performed a baseline, monthly, and final 6MWT on a suitable walking path determined by a visit to the participant’s home. The monthly 6MWTs were unsupervised and their start time was recorded with a dedicated button on the device. Step count and walking speed correlated to 6MWD. They concluded that tele-accelerometry is feasible and it provides a useful indicator of exercise capacity. They also showed that the remote 6MWTs were safe in HF patients. This is one of the only studies that focused on the practical application of wearable activity in telemedicine.

Dontje et al (2014) gave 68 patients a tracking armband for 2 consecutive weekdays to understand activity levels in HF patients compared to guidelines. 44% of patients were active for less than 30 minutes per day, which was labeled as sedentary. Higher NYHA classification was also associated with less daily activity. Psychological characteristics, like self-efficacy and motivation (measured through questionnaires), were also linked to activity levels. This suggested that healthcare providers should promote activity in sedentary patients by delivering advice aimed at improving motivation and self-efficacy.

Izawa et al studied the relation between physical activity and exercise capacity with the goal of determining target levels of physical activity for improving exercise capacity. They used an electronic pedometer to measure the steps 24 hours per day for 1 week in 97 middle-aged and 60 elderly patients. Using receiver operating characteristic curves, they found that approximately
6000 steps per day was the cut-off for an exercise capacity threshold of 5 metabolic equivalents (METs). Patients below 5 METs are considered to be at high risk of death and have significant difficulty performing activities of daily living.\textsuperscript{49}

Conraads et al (2014) pooled data from 781 HF patients with ICDs with continuous accelerometer sensors.\textsuperscript{19} This was an expansion on a similar earlier study by Braunschweig et al.\textsuperscript{50} The sensor in the devices stored the number of active minutes per day for up to 425 days and was collected during clinical visits. Physical activity in the first 30 days of monitoring strongly predicted hospitalization or death. A multivariable analysis was significantly improved when incorporating physical activity, and there was a 4\% reduction in risk for each 10 minutes per day additional activity. Even a relatively short period daily activity monitoring can predict outcomes that occur over one year later. Note that these sensors are included in implants to help the devices detect changes in activity levels so it can adjust the heart rate accordingly. While the use of the data from these sensors is intriguing, not all patients have these implants, and these implants can only transmit to their companion devices. This is only done during clinical visits.

Recently, Brooks et al (2015) showed that a smartphone application for self-administered 6MWT yielded accurate repeatable measurements both in the clinic and at home.\textsuperscript{34} After initial phases to calibrate and test, they gave the app to 19 patients (including both HF and pulmonary hypertension patients) to perform 6MWTs at home several times per week for 2 weeks. They also showed that patients would be willing and able to perform frequent self-administered tests at home using the application.
<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Participants</th>
<th>Technology used</th>
<th>Duration</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davies (1992)</td>
<td>20 normal, 30 stable HF, 20 age-matched controls</td>
<td>Mercury switch pedometers on wrist and ankle</td>
<td>Up to 14 days</td>
<td>Correlated with questionnaires and exercise tests</td>
</tr>
<tr>
<td>Oka (1993)</td>
<td>45</td>
<td>Device on belt with HR and arm and leg motion sensors</td>
<td>2 days</td>
<td>Daily activity lower than exercise capacity</td>
</tr>
<tr>
<td>Walsh (1997)</td>
<td>84</td>
<td>Mechanical pedometers, 1 on each hip,</td>
<td>1 week, then followed for outcomes</td>
<td>Daily activity predicts death and may outperform laboratory tests</td>
</tr>
<tr>
<td>Jehn (2009)</td>
<td>50</td>
<td>Accelerometer, belt on left hip, data copied from device on return</td>
<td>7 days</td>
<td>TWT correlated with VO₂ and 6MWD, time spent fast walking strong classifier for NYHA</td>
</tr>
<tr>
<td>Jehn (2013)</td>
<td>155</td>
<td>Accelerometer, transmitted data via mobile network, belt on hip</td>
<td>Monthly 6MWT for 1 year</td>
<td>Tele accelerometry feasible and a good indicator of exercise capacity</td>
</tr>
<tr>
<td>Dontje (2014)</td>
<td>68</td>
<td>Armband with accelerometer, upper arm</td>
<td>2 consecutive weekdays</td>
<td>Many patients active &lt;30 min/day, higher NYHA associated with lower activity</td>
</tr>
<tr>
<td>Izawa (2012)</td>
<td>157 (97 middle-aged, 60 over 65)</td>
<td>Belt</td>
<td>1 week</td>
<td>6000 steps/day is the cut-off for 5 METs</td>
</tr>
<tr>
<td>Conraads (2014)</td>
<td>781 (pooled data)</td>
<td>Sensor in ICDs</td>
<td>Up to 425 days of data, focused on first 30 days</td>
<td>First 30 days of daily activity predicted outcomes</td>
</tr>
<tr>
<td>Brooks (2015)</td>
<td>19 (for home phase)</td>
<td>iPhone in pocket</td>
<td>2 weeks (~3 tests per week)</td>
<td>Easy to use with accurate and repeatable results</td>
</tr>
</tbody>
</table>
4 Objectives and hypothesis

4.1 Gap in knowledge
Unlike supervised exercise testing, daily activity is unpredictable and uncontrolled. One might expect that the extraneous variables involved in tracking activity throughout a patient’s daily life would drown out any useful information about a patient’s HF. However, the review of literature in Chapter 3 has revealed numerous studies using different tools and approaches to show that daily activity monitoring is useful in HF. Some have even argued that it has the potential to outperform clinical exercise tests as a predictor of mortality and hospitalization. Given this, and the fact that this has been studied for over two decades, why do current telemonitoring interventions still not incorporate activity monitoring? This may be because of the limitations of the technology used in previous studies, and the complexity of transferring those findings to academic practice. Previous studies typically used devices worn on the hip or upper arm, and often imposed requirements on patients to wear the device for a minimum amount of time each day or to perform certain activities on a regular basis. A clinical telemonitoring intervention requires seamless and reliable data transfer, both patient and clinician cooperation and motivation, and should not be so complex that it interferes with existing workflows. Wristbands, while potentially less accurate, are more convenient and comfortable than an armband, and are not attached to clothing like a device worn on the hip. Consumer wristband activity trackers are relatively inexpensive, easy to acquire, and are designed to be easy to use. They have built-in data wireless synchronization with mobile phones and can last several days before being needing a recharge. While a consumer wristband activity tracker provides many advantages over previously used research tools, two major questions exist:

- Are these consumer devices reliable enough to provide data that can be clinically relevant?
- Can step data from these consumer devices be related to clinical measurements of exercise capacity?

Consumer devices are subject to lower standards than the kinds of tools typically chosen by researchers. Furthermore, although wristband devices are reasonably accurate, they do count extra steps from hand movements that do not result from walking. However, the purpose of the
activity tracker in this context is not to achieve an exact count of the number of steps taken, but to evaluate the wearer’s activity levels. Clinicians are also accustomed to the exercise testing tools they use to evaluate HF. If step data from a consumer device provide an indication of HF severity, it should also correlate with results from clinical measures of exercise capacity.

The study was originally aimed at correlating the proportion of sedentary behaviour to HF severity. However this approach was problematic for three reasons; 1) the tools used could not reliably distinguish between inactivity due to lack of movement and not wearing the activity tracker, 2) a sedentary lifestyle could be, at least in part, based on occupation and other factors rather than heart failure, and 3) symptoms of heart failure are observed when a patient is performing near peak activity, and the clinical tests actually evaluate maximal exercise. Focusing on peak activity levels revealed stronger correlations with CPS data, and patients did not need to wear the devices every possible minute for the analysis to be valid.

4.2 Hypothesis

The purpose of the present work was to develop a baseline for the relationship between heart failure and step counts obtained from consumer wristband devices, via clinical tests of exercise tolerance. A patient’s condition is not expected to change significantly in only two weeks, assuming no decompensation events. Therefore it is reasonable to expect that the exercise capacity results from the exercise test are relatable to the daily activity immediately following the test. The hypothesis is that lower exercise capacity results in the CPS would show correspondingly low results in the wristband step data. To test this hypothesis, we provided consumer activity trackers to HF patients to collect activity data during the weeks following a CPS. At the same time, this study would show whether these devices are usable and that their built-in data collection capacities are suitably reliable. There are many different ways raw minute-by-minute step data could be related to the various outputs from a CPS. The analysis of the results required comparison and optimization of metrics derived from step data to determine simple but effective indicators for HF that can be calculated from raw step data.
5 Activity monitor study

5.1 Methodology

5.1.1 Planning and device testing
Preparation for this study involved evaluating several different exercise monitors based on several different requirements and criteria, these included:

Requirements:
- Wristband
- Battery must last longer than 1 day
- Must be readily available at the consumer-level
- Data export: step data must be accessible and exportable for analysis
- Simple or no display

Criteria:
- Price
- Ease of use: wearing, charging, data collection
- Durable and water resistant
- Unobtrusive: device should not be too colourful, should not make excessive noise or light up too much, and should not interfere significantly with daily activities
- Battery life
- Storage (i.e. how long the device can go without synchronizing with an external device)
- Device warranty

We researched many different options, and physically tested the following devices:
- Fitbit: Flex, Force, Charge HR
- Jawbone: UP, UP24
- Withings Pulse

Some of the specifications of these devices had to be confirmed by contacting their customer support. It was challenging to get a clear response from some of them, especially with respect to questions about how long the device could go without synchronizing. This was presumably because these devices were not designed to go very long without syncing. The Fitbit Flex was
one of the only products where this was explicitly stated in publically posted materials. The Fitbit Flex can go a full week without synchronizing before it starts to lose minute-by-minute data, and can go 30 days before it starts to lose daily total steps.

The Fitbit Flex meets the criteria in the following ways:

- Inexpensive (approximately $100)
- Replaceable wristband (i.e. if participant needed a smaller size, or a wristband was damaged it could be easily replaced)
- Display does not show actual number of steps, just dots representing percentage of daily goal accomplished
- Battery life up to 5 days, can go without syncing for up to a week (or up to a month but with only daily summaries)
- Easy to use, easy to set up, syncs reliably
- Data is easily accessible through the Fitbit application program interface (API)
- Water resistant
- Compatible with a wide range of mobile devices and computers
- Responsive customer support

Of course, the Fitbit Flex has some disadvantages (this list excludes the technical issues observed during the study period), which include:

- There is no way to distinguish between inactivity and not wearing the device
- One cannot check battery level on wristband (except when charging). Battery level can only be checked on the mobile app, or online after a recent sync
- The tracker has a sleep-tracking feature that can be accidentally activated during certain activities, such as while riding a bike or pushing a shopping cart. There is no way to disable this feature. This does not affect step data but it can confuse participants.
- Clasp to hold the wristband onto the wrist can be quite difficult to close with one hand, especially for first-time users or people with limited dexterity
- The tracker has to be removed from the wristband and inserted into the charger in a specific orientation for charging; proper re-insertion of the tracker is also difficult and non-intuitive
- Can experience charging issues after continued use over a long period of time, which
can be due to build-up on the exposed charging contacts.

Before giving devices to any patients, the materials were tested while mimicking the study protocol, specifically to make sure the tracker continues to synchronize with the phone after several days even when the phone remains locked, after the phone restarts, or after the tracker runs out of battery. The materials and instructions were given to two colleagues for testing and feedback on the methodology and usability.

5.1.2 Recruitment

Planning and ethics approval took place during the spring of 2014. UHN Research Ethics Board approval for the study was granted in August 2015. Participants were recruited between September 2014 and June 2015. Participants were recruited from a cardiac clinic at TGH. Patients were recruited based on the following inclusion and exclusion criteria:

**Inclusion criteria:**
- At least 18 years of age
- Stable NYHA Class II or III
- Left ventricular ejection fraction $\leq 35\%$
- Normal walking without walking aids
- Capable of understanding instructions in English
- Ability to wear, care for, charge, and return the study materials
- CPS prior to recruitment (all but 2 participants completed their CPS the same day as the recruitment)

**Exclusion criteria:**
- Congenital heart failure
- HF diagnosis less than 6 months prior to recruitment
- Traveling out of Canada for more than 1 week during the study period (data plans on the phones were from a local cellular provider)

Daily recruitment rates ranged from 0 to 3 participants. The TGH clinic was open Monday through Wednesday, with several periods where the clinic was closed or there were no eligible patients. A recruiter was not always available during clinic hours.

The recruiter brought the study kits to the clinic during recruitment and potentially eligible
participants were identified based on the CPS appointments for that day. The recruiter first checked with the primary doctor or fellow for an up-to-date assessment of eligibility. Then the recruiter would approach the participant, introduce the study, demonstrate the use of the study materials, go over the consent form (Appendix B), get signed consent, and then collect health and contact details (Appendix C). Participants were compensated $50 for successful completion of the study and return of the study materials.

5.1.3 Materials
Participants received a study kit (Figure 2) containing a Moto-G Android smartphone, phone charger, Fitbit Flex activity tracker, Fitbit Flex wristband, Fitbit Flex charger, dual USB-port wall adapter (so that both the phone and the Fitbit could be charged simultaneously), and a FedEx box with a prepaid shipping label to return the materials.

Figure 2: Study kit components (phone, Fitbit Flex, charging accessories)

Participants were also given a guide with instructions for proper use of the kit components and information about the study (Appendix D).
All phones were connected to a mobile data plan to ensure continuous data synchronization and so the participant did not have to interact with the phone or have a wireless network. The phones were locked with a security code that was not shared with the participants to ensure they did not accidentally modify any of the settings. Participants were still able to turn the phone off via a hardware button, or disable its wireless signals from the lock screen. The Fitbit Android application was installed on each phone and the application was paired with the kit’s Fitbit device. The phones and the activity trackers were charged and the connection was verified before giving them to a participant.

Kit materials were labeled with numbers to ensure the phones and the activity trackers were not mixed up. The phones were locked with a code that was not shared with the participant to prevent them from accidentally modifying the settings on the phone or using the phone for non-study purposes. They were asked a series of questions (Appendix C) about their HF history, weight, age, height, sex, handedness, wrist preference for the Fitbit, and regular physical activities. Profiles for the Fitbit accounts were also updated with each participant’s weight, height, sex, and age.

Ten kits were used throughout the study. Participants were asked to return the study materials soon after their study period ended. Kits were inspected, cleaned, recharged, and then re-used for new participants.

5.1.4 Cardiopulmonary study
Each participant undertook a clinically indicated CPS on a Lode Corival (Netherlands) cycle ergometer with a 10-watt step protocol. The clinic’s specialist conducted the CPS without any involvement from the recruiter or any study-related deviation from the normal protocol used in that clinic. The patient terminates the test when they no longer feel comfortable enough to continue, usually due to leg fatigue. Patients are monitored with a 12-lead ECG, a neoprene facemask to collect respiratory gases for analysis by a Medgraphics Ultima CardiO2 (Minneapolis, MN), pulse oximetry, and periodic blood pressure measurements. During recovery, patients pedal for 2 min at 10 watts, followed by 3 minutes of seated rest. The test output includes absolute and relative VO$_2$ peak, peak watts on bicycle, as well as predicted and percent of predicted values based on independent factors such as age and sex. CPS summary
reports were obtained with the patient’s permission from the clinic specialist (an example is shown in Appendix A).

5.1.5 Remote activity assessment

Participants took the kit home and were instructed to plug the phone in to charge for the duration of the study in a room they access regularly, and to wear the Fitbit on their wrist as much as possible for the duration of the study period. The study period was targeted at two weeks starting the day after recruitment, although some study periods were extended because of technical issues or personal reasons. Some participants took the phone with them when they were away from home for more than a day, such as a visit to their cottage. As long as the phone was connected to the Internet (through cellular data), connected to the Fitbit Flex (via Bluetooth), and in range (approximately 10 meters, or less if there are obstructions or sources of interference), it would synchronize data every 20 minutes. The Fitbit Flex tracker can go several days without synchronizing with the phone without any data loss, but participants were encouraged to be in range of the phone at least daily. The Fitbit tracker has a battery life of five to seven days, and participants were asked to charge the Fitbit at least every three days, and preferably to charge it while sleeping. Participants were also asked not to wear the Fitbit Flex while showering or swimming.

The targeted two-week study duration was selected to give participants enough time to get used to the wristband as well as to include both weekdays and weekends in the collection period to allow a full picture of their physical activity. We also assumed that a participant’s functional status would not change significantly within such a short period, and thus all days in the study were treated as comparable to the CPS done immediately before the start of the study. The targeted two-week duration also helped ensure that even if a few days were missed for technical or personal issues, there would still be several days of data that could be used.

Anonymous daily data was synced regularly in an online Google Sheet document with a custom Google Script so that all the step data, syncing, and battery levels of the kits could be downloaded and monitored simultaneously. Upon completion of the study the intraday data was downloaded to a similar Google Sheet with a slightly different script and saved for analysis. These scripts were customized based on open source software that use Google’s libraries to
connect to the Fitbit API and download data from each Fitbit account used in the study. An application was created on the Fitbit developer website, which provided secure keys to use for authorization. Each of the 10 Fitbit accounts was set up to give permission to access by this application, which resided in the Google Script. Modified public versions of the Google scripts can be found here: https://github.com/simonbromberg/golefitbit.

5.1.6 Data analysis
Data was analyzed using Microsoft Excel, SPSS, and R Studio. The Fitbit API returned number of steps and the date and time for every minute from the beginning to the end of each individual study period. This data had to be processed for analysis. Some processing, such as simple sorting, averaging, and counting, could be done quite easily in a spreadsheet. However more complicated methods required a different approach. A Mac OS X application written in Swift (selected mainly due to convenience and personal preference) was created to run custom algorithms for processing raw step data. These algorithms summarized the raw data, returning results such as the best daily 10 minutes or the longest or best daily walk. The application parsed aggregated data stored in a single comma separated value file. Rows with zero steps were removed to improve performance. Multiple variations for individual metrics can be quickly outputted for all participants, such as best 3, 6, 10, 30, and 60 minute walks. These outputs were correlated with CPS results and NYHA classification using linear or logistic regression, respectively.

5.2 Results
5.2.1 Study population
51 eligible participants were recruited, with one participant leaving the study before any data was collected for unrelated reasons. One other patient was recruited but then soon afterwards deemed to be ineligible because they had congenital HF. One participant performed a treadmill exercise test instead of a bicycle ergometer, so their CPS results did not include watts. This participant was female and classified as NYHA II. The study population and mean exercise data is summarized in Table 2 below. The majority of participants were men, over 50, and in NYHA class II. Means in Table 2 are reported with \( \pm \) one standard deviation, and showing p values where the difference in means between NYHA II and III is statistically significant, calculated
with a one-tail t-test. Several of the differences in the group means are statistically significant, including step data outputs. Although NYHA classification is subjective, and the distinguishing between class II and III may not be easy, there are clear differences between the two groups with respect to exercise capacity.

Table 2: Study population and exercise data summary
Statistically significant p-values according to a one-tail t-test are indicated where applicable

<table>
<thead>
<tr>
<th>Cohort</th>
<th>NYHA II</th>
<th>NYHA III</th>
<th>Total</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>35</td>
<td>15</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>29</td>
<td>14</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>54 ± 14</td>
<td>56 ± 12</td>
<td>54 ± 13</td>
<td></td>
</tr>
<tr>
<td>Height (cm)</td>
<td>173 ± 9</td>
<td>175 ± 8</td>
<td>174 ± 8</td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>85.5 ± 20.6</td>
<td>92.8 ± 18.3</td>
<td>87.7 ± 20.0</td>
<td></td>
</tr>
</tbody>
</table>

| CPS Results                  |         |         |       |      |
| Relative VO\(_2\) peak mL/(kg·min) | 14.9 ± 3.6 | 12.9 ± 3.6 | 14.3 ± 3.7 | < .05 |
| % of predicted relative VO\(_2\) peak | 51.4 ± 14.2 | 43.8 ± 9.7 | 49.1 ± 13.4 | < .05 |
| Absolute VO\(_2\) peak (L/min)   | 1.27 ± 0.43 | 1.19 ± 0.45 | 1.25 ± 0.44 |      |
| % of predicted absolute VO\(_2\) peak | 57.7 ± 17.1 | 50.5 ± 11.6 | 55.5 ± 15.9 | < .05 |
| Watts (n = 49)               | 96 ± 31  | 77 ± 33  | 90 ± 33 | < .05 |
| % of predicted watts (n = 49) | 63 ± 22  | 48 ± 14  | 59 ± 21 | < .005 |
| CPS duration (minutes)       | 10.1 ± 3.0 | 8.1 ± 3.3 | 9.5 ± 3.2 | < .05 |

| Daily Activity              |         |         |       |      |
| Average daily steps         | 5864 ± 2273 | 4301 ± 2291 | 5396 ± 2368 | < .05 |
| Average daily walking time (minutes) | 206 ± 65  | 171 ± 72  | 196 ± 69 | < .05 |
| Most steps in 10 minutes    | 1016 ± 186 | 772 ± 190 | 943 ± 217 | < .001 |
| 99\(^{th}\) percentile steps in 1 minute | 105 ± 12  | 92 ± 11  | 101 ± 13 | < .001 |

5.2.2 Cardiopulmonary study results
As one would expect, exercise test results correlated strongly with each other, such as VO\(_2\) peak versus watts (Figure 3). Percent of predicted results also correlated strongly with each other such
as percent of predicted \( VO_2 \) peak versus percent of predicted watts (Figure 4). Correlations between absolute values and predicted values were slightly weaker (Figure 5). Watts achieved during the CPS show the integrated response of the body to exercise. However, while the literature is clear on the predictive power of exercise capacity, there is no discussion of the value of this simpler output from the test. Thus it was worth showing that the watts achieved during the CPS correlates not only with percent of predicted exercise capacity, but raw exercise capacity as well.

![Graph showing linear correlation of absolute \( VO_2 \) peak versus watts.]

Figure 3: Linear correlation of absolute \( VO_2 \) peak versus watts

- \( r = 0.89 \)
- \( n = 49 \)
- \( p < .001 \)
Predicted values were useful ways to correct for factors like age and sex so that a wide range of different kinds of patients could be correlated. For example, even a healthy man in his seventies may have a lower relative VO\textsubscript{2} max than a young man with HF. No trend was observed between
age and VO₂ peak (Figure 6) but the participants included have different HF severities that is often but not necessarily dependent on age. Figure 7 shows that age does not appear to be a dominant factor in functional status (as indicated by NYHA classification) within the study population.

**Figure 6: Age versus relative VO₂ peak**

**Figure 7: NYHA versus age box plot**
5.2.3 Estimation of peak activity from step data

Raw step data was collected as 1440 entries per day, showing estimated steps for each minute from midnight to midnight. This corresponds to approximately 20,000 rows per participant. The majority of these data were zeros, or no steps taken. The average total walking time per day ranged from under 1 hour to almost 6 hours, or 60 to 360 minutes. This represents 4–25% of the day, or 8–50% of the waking day. The median of average walking times was 3.4 hours. Not all activity was captured, since wearing adherence varied between participants. Therefore, rather than look at total daily activity, which includes ambiguous periods of zero activity, we should focus instead on minutes with activity.

Steps taken in one minute (ignoring zeros) ranged from 1 to 180, with a median of 19. The lowest individual maximum steps taken in one minute was 98, with median maximum steps of 118. The distribution of steps can be visualized by graphing the number of steps in each active minute against each point’s fractile (Figure 8). Similar distributions were observed for all participants. The interesting features of these distributions are that the majority of activity is below 60 steps per minute, and that there is significant spread for peak activity levels. This indicates that simply looking at the minute with the most steps for each participant is not a good representation of peak activity, because those values are outliers. Averaging the steps per minute would also hide peak activity, since the majority of the data is far below peak activity. Thus approaches like using upper percentiles, averaging the top 10 values, or grouping consecutive minutes together are more robust statistics for estimating peak activity and less sensitive to noise. These measures of peak activity are based on all of the activity data over the course of the two-week study period for each participant.
5.2.4 Step and CPS data correlation

Many different approaches to using raw minute-by-minute step data were evaluated to find the optimum correlations with exercise test data. Metrics considered included average or maximum steps per day, \( n^{th} \) percentile step minute, most steps in \( x \) minutes, longest walk, and total walking time. Some of these approaches are summarized in Table 2, showing statistically significant differences between NYHA classes. The quality of each metric was evaluated based on its correlation coefficient and statistical significance with CPS data results. The CPS summary reports included up to 36 different values for each participant (see Appendix A). Eight of these were for measurements taken at rest, and the remainder for values related to the end of the CPS or for predicted results. While each one of these could be correlated to step data, VO\(_2\) is of particular interest because it is the strongest indicator of exercise capacity. Peak resistance achieved on the stationary bike also provided an indication of the functional capacity of the individual, representing the combined response of the body’s systems to drive muscles to perform work at peak levels. The best CPS columns for correlating to proposed step data metrics were also verified through a correlation matrix with Pearson coefficients and statistical significance.

Metrics corresponding to maximal or near maximal walking activity yielded the best
correlations with CPS results. This makes sense because the CPS is aimed at measuring peak exercise performance. Correlations of step data with raw CPS values (e.g. Figure 9) were weaker than correlations based on percentage of predicted values. Predicted values help correct for independent factors like age and sex that the current study cannot account for due to its limited sample size. Best 10 minutes and 99th percentile steps per minute correlated best with CPS data, particularly percent of predicted watts (Figures 10 and 11). These correlations are also summarized in
Table 3. Duration of the exercise test did not correlate as strongly with step data (Figure 12). Average maximum daily steps also did not correlate as strongly for percentage of predicted relative VO\textsubscript{2} peak (Figure 13) but did for percentage of predicted watts (Figure 14). An alternative to 99\textsuperscript{th} percentile steps is the average of the top ten minutes by number of steps (Figure 15).

**Figure 9: Relative VO\textsubscript{2} versus most steps in 10 minutes**
Table 3: Correlation of predicted values and step metrics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Predicted watts vs Best 10 minutes, 99\textsuperscript{th} percentile steps</td>
<td>0.57</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>% of Predicted relative VO\textsubscript{2} peak vs Best 30 minutes</td>
<td>0.49</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>% of Predicted absolute VO\textsubscript{2} peak vs Best 10 minutes</td>
<td>0.43</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>% of Predicted watts vs maximum daily steps in 1 minute</td>
<td>0.54</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

![Figure 10: % predicted watts versus 99\textsuperscript{th} percentile steps in 1 minute](image)

$r = 0.57$
$n = 49$
p < .001
Figure 11: % predicted watts versus most steps in 10 minutes

Figure 12: Exercise test duration versus 99th percentile steps
Figure 13: % predicted relative VO$_2$ peak versus average maximum daily steps in 1 minute

Figure 14: % predicted watts versus average maximum daily steps in 1 minute
Another approach that was tested was best overall walk, both by duration and number of steps, with the end of a walk being defined by when the time until the next non-zero step exceeded a certain gap size. This gap size could be varied according to what would be considered a reasonable resting time. One could also consider walks that are only at or above a certain pace. There are a large number of parameters one could adjust to optimize the result for any correlation, although the choice of these parameters would have to be justifiable. However, these more elaborate approaches did not beat the simpler methods described above (e.g. Figure 16).

The best walk of a participant may be too heavily influenced by personality and lifestyle to provide useful representation of exercise capacity. Some may prefer long slow walks, while others prefer quick but short walks, and some may not have any time or interest in walks at all. However, when considering short walking periods that are consistent across all participants, such as individual minutes or ten minutes out of hundreds or thousands of minutes with activity, there are bound to be multiple cases where the walking is on the high end of each individual’s capacity.

**Figure 15**: % predicted relative VO\textsubscript{2} peak versus average top 10 steps in 1 minute

\[ r = 0.37 \\ n = 50 \\ p < .005 \]
Figure 16: % predicted relative VO$_2$ peak versus steps in longest walk (maximum gap 2 minutes)

Proportion of active time spent above a certain pace also did not correlate with exercise test results. Figure 17 shows an example of the relationship between the percentage of active time spent walking faster than 60 steps per minute and percent of predicted relative VO$_2$ peak.

Figure 17: % predicted relative VO$_2$ peak versus % of active minutes at >60 steps/minute
Step data metrics based on daily totals also did not correlate as well with exercise test results, presumably because the devices could not identify whether zero steps at any particular minute was because the person was not moving, or because they were not wearing the device. These metrics involved walking time, or daily total steps. While slightly weaker than other correlations noted above, average daily walking (Figure 18) and average daily total steps (Figure 19) did show potential.

Figure 18: % predicted relative VO$_2$ peak versus average daily walking time
Stronger correlations were observed within age-based subsets of the data, particularly for participants 60 years of age and older (Table 4, Figures 20 and 21).

**Table 4: Correlations within subgroup where age ≥ 60**

<table>
<thead>
<tr>
<th>Correlation</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>% predicted watts versus 99&lt;sup&gt;th&lt;/sup&gt; percentile step per minute</td>
<td>0.76</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Relative VO&lt;sub&gt;2&lt;/sub&gt; peak versus best 6 minutes, best 10 minutes, or 99&lt;sup&gt;th&lt;/sup&gt; percentile steps per minute</td>
<td>0.66</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Absolute VO&lt;sub&gt;2&lt;/sub&gt; peak versus average minutes of walking per day</td>
<td>0.62</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>% predicted relative VO&lt;sub&gt;2&lt;/sub&gt; peak versus 99&lt;sup&gt;th&lt;/sup&gt; percentile steps per minute</td>
<td>0.74</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

The other subgroup (age < 60) also yielded some significant correlations, such as % predicted watts versus average maximum daily steps in 1 minute ($r = 0.61$, $p < .001$). The subgroup consisting of participants with ages below 40 showed correlation coefficients as high as 0.85, but this only included 10 participants.
5.2.5 NYHA classification prediction
As shown in Table 2 above, there are significant differences between the two NYHA classification groups, especially with respect to the step data metrics (Figures 22 and 23). The
full range of NYHA II actually appears to be more spread out and overlap somewhat with participants in NYHA III, but the more severe HF patients are concentrated at the lower end. This is probably a consequence of the subjectivity of NYHA classification. The difference between II and III is a subtle one, but the more severe a patient’s condition is, the more likely they will be assigned a higher classification. Using logistic regression the 99th percentile steps in 1 minute versus NYHA classification generated a significant model (p < .001). The receiver operating characteristic (ROC) curve is shown in Figure 24 and has an area under the curve (AUC) of 0.7924.

![Figure 22: Most steps in 10 minutes versus % predicted watts showing NYHA by color](image_url)
Figure 23: NYHA classification versus 99th percentile step minute box and whisker plot

Figure 24: ROC curve for classifier based on 99th percentile steps in 1 minute, AUC = 0.7924
Average of proportional time spent walking above a certain pace also showed statistically significant differences between NYHA groups (Figure 25). This compares the percentage of time, out of the total observed active minutes, where the participant walked at a pace greater than 60 steps per minute. This analysis was an attempt to reproduce a result in a paper from Jahn et al, where the proportion of total wearing time spent walking at a fast pace (83 to 115 m/min) correlated strongly with NYHA classification. Using a threshold greater than 60 steps per minute gave very similar results. Also this analysis is based on total active time, not total wearing time, as the latter is not directly available.

![Box plot of NYHA versus % of active minutes where > 60 steps/minute (p < .05)](image)

**Figure 25: Box plot of NYHA versus % of active minutes where > 60 steps/minute (p < .05)**

### 5.2.6 Technical evaluation

Since part of the purpose of the study was to study the feasibility of using these consumer devices as a remote monitoring tool, we kept track of any technical or usability problems that were observed during data collection.

Problems experienced during data collection either resulted from technical problems with the device or because of individual situations. 35 participants did not experience any issues with data collection. Nine participants had resolvable problems that did not interfere significantly with data collection. Two participants had some unresolved technical issues that resulted in somewhat less than 14 days of data, although these still had several days worth of usable data. Four participants...
had personal situations that resulted in less data collected.

Issues that were easily resolved by troubleshooting via email or phone included:

- 3 participants’ trackers stopped automatically syncing for unknown reason (resolved by giving participant temporary access to the phone and having them re-launch the Fitbit app to force a synchronization),
- 2 participants forgot to leave the phone plugged in and at home,
- 1 participant’s Fitbit was not paired properly and had to give participant access and instructions to pair it again,
- 1 participant paired the Fitbit with their personal phone not realizing that would interfere with the study,
- 1 participant put phone into airplane mode,
- 1 participant forgot to charge the Fitbit

Issues interfering with data collection included:

- 2 delays in synchronization causing some loss of intraday data,
- 1 Fitbit battery issue,
- 1 participant wore the wristband infrequently,
- 1 participant experienced unrelated health problems
- 1 unexplained issue

However, all participants had at least six days of with greater than 100 steps, and at least 12 hours of aggregated walking time.

The pairing of the Fitbit Flex with the smartphone was straightforward and generally reliable. Data synced regularly and automatically in the background and most participants had no trouble charging and wearing the device. However, in several cases data stopped synchronizing from the paired Fitbit. This may have been just because the participant was not in range that day. If a participant had several days without incoming data, or several days with no activity, they would be contacted to troubleshoot the problem. In most cases, the participant had forgotten to leave the phone plugged in, the phone was in airplane mode, or the Fitbit was not charged. However, in some cases the Fitbit just would not sync, despite being fully charged and in range of the phone with a data signal. This was resolved in some cases by sending the phone’s unlock code to the participant so they could open the Fitbit app to force it to connect to the tracker. After this
happened I started unlocking the phones and checking the connection before giving the kits to the participant. In a couple of cases, the Fitbit would not sync reliably, would not charge, or the battery was draining rapidly, and I simply contact Fitbit and they sent me a replacement under warranty.

The Fitbit API provides both daily totals and intraday data. In some cases, the daily total according to the sum of the intraday data did not exactly equal the daily total according to Fitbit. These differences were mostly very small. However, if the device goes more than seven days without syncing, it will start to lose intraday data. Participants that were unable to sync their device wound up losing several days of data. In most cases, once syncing was restored, the missing days were filled in, but one participant lost nine days of intraday data (leaving six days with intraday data) in the middle of their study period because of delayed syncing.

The rubber wristbands were somewhat prone to damage, but Fitbit would replace these upon request. Most participants used the large-size wristbands that came with the devices, but one participant could not be recruited because their wrist was too big. Fitbit has extra-large wristbands available (no charge) upon request. Any future study should consider having one or two of these larger wristbands on hand, or if using a different device, ensure it can accommodate different wrist sizes. Two potential participants were not recruited because they work a lot with their hands and were worried they would damage the Fitbit Flex.

Participants also consistently returned the kits without losing or damaging any components. Some participants held on to the kits for several weeks longer than the study period, and they had to be contacted to remind them to return the kits. Eight participants took almost two weeks to send the kit back, and two participants took over one month. However, this was an expected limitation of the study design and did not cause any significant delays in the recruitment process because there were sufficient extra kits.

Access to Fitbit data through the Fitbit API was simple to set up, however, on a couple of occasions the script stopped working. The first time it started working again shortly afterwards. The second time was because Google’s OAuth1.0 library had been deprecated. By then Fitbit had released its OAuth2.0 beta, so the problem was resolved by redesigning the script so that it could connect to the API via OAuth2.0 using Google’s OAuth2.0 library.
6 Discussion

The present work has identified several new step data based metrics that can be used to relate patient daily activity to exercise capacity, and thus provide a remote, continuous indicator of the severity of their condition. The study also provided evidence of the usability and reliability of consumer activity trackers for remote monitoring of physical activity in HF patients.

6.1 Step data correlations

Apart from the strong correlations of the elderly subgroup, the largest linear correlation coefficient for the full study population was approximately 0.6. This is close to the result from a linear regression in the study by Jehn et al, who reported a correlation coefficient between TWT and VO$_2$ peak of 0.72. Izawa et al correlated exercise capacity to daily activity by finding the cut-off number of daily steps for 5 METs based on logistic regression. Many of the studies focused on prediction of NYHA classification from daily activity data or did not attempt to correlate exercise capacity to daily activity but instead sought to better understand daily activity levels of HF patients and relate them to outcomes directly. By first showing that consumer devices can collect data that is relevant to exercise capacity, we can justifiably use these devices in longer-term studies to examine outcomes or compare daily activity to other indicators of HF, such as weight and blood pressure. The slightly higher correlations in the Jehn paper may be due to the use of a device worn on the belt that provided accelerometer data so wearing time could be detected, or different recruitment methodologies. Removing outlying values improved the strength of some of the correlations slightly, but no participants were identified as consistent outliers. Outperforming previous studies, which used activity trackers designed for research, using a consumer device was not the intention of the present work. Devices designed for research would be expected to provide better results, because they provide more detailed data with greater storage capacity so data would not have to be synchronized until the end of a study. Rather, this study aimed to show that consumer devices performed reasonably well by correlating the data from these devices with clinical tests.

During a 6MWT patients are intentionally walking at a pace that approaches their capacity, whereas during their normal daily walking, they would not necessarily push themselves as much. However, the results show that when collecting continuous step data over several days, there are
enough instances of activity approaching capacity to correlate with CPS data.

Previous similar studies reported findings based on frequency-based measures of daily activity, such as mean steps per day\textsuperscript{52} and total walking time per day\textsuperscript{47}. These analyses rely on participant adherence with regard to wearing the activity monitor, or the ability to identify periods where the participant is not wearing the device. If participants were not required to wear the tracker a certain amount of time per day, a participant’s daily total may be lower, not because of a decreased exercise capacity, but because they wore the device less. Averaging daily totals also makes very different activity patterns look the same. When averaged, walking 5000 steps every day of the week would appear equivalent to walking 1000 steps on 3 days and 8000 steps on 4 days. By contrast, the most steps in 10 minutes or the 99\textsuperscript{th} percentile does not depend at all on how often the device was worn, just as long as a sufficient sample is considered so that it includes activity that approaches peak levels. On any given day, an individual may not walk very much, but it is reasonable to expect that at least part of the time they will be walking at a walking pace that is a good representation of their functional capacity.

The best performing step data metrics, 99\textsuperscript{th} percentile steps in 1 minute, and best 10 minutes, represent peak activity without being thrown off by outliers. Maximum steps in individual minutes can show a large number of steps (as high as 180 steps in this study) that may not be an accurate representation of typical peak exercise. These step metrics are easy to calculate and can be calculated over almost any time period. In longer-term studies or in clinical practice they could be calculated on a rolling week basis, to show trends in patient status.

6.2 Advantages of the study design

Recruitment of HF patients was fairly unrestricted, as long as patients fit within the target NYHA classes and did not have unstable conditions or walking problems, they would be included. Some of the other studies in this area had similar open recruitment strategies.\textsuperscript{47} In a practical application, clinicians must be able to apply the remote monitoring tools at their disposal to any patient that might benefit from it, not only participants fitting specific criteria. This study also used the tracking devices in a fashion that would be similar to how they would be used as a practical remote monitoring tool, rather than focusing entirely on data collection and strict controls on how patients went about their daily lives.
By giving participants kits very similar to what they would use if this were an actual clinical tool, this study has provided a good representation of the benefits as well as the potential pitfalls of a telemonitoring system that relies on these consumer devices. Previous studies performing similar interventions generally were not attempting to simulate how these concepts would work in a telemonitoring application. Jehn et al did build on their previous work by using a tele-accelerometry system, but this was not a wristband consumer device, but rather a research oriented accelerometer sensor worn on the belt.48

Although there were technical issues, many resulted from the way the study was designed, since it required participants to use the devices in a way that deviated from their intended use. In a longer study, participants would have more time to get used to the devices and be given more responsibility over their care. They might even pair them with their personal phones, and would have access to the companion mobile application to view their incoming data and to make sure it was still connected.

6.3 Study limitations

6.3.1 Study population
This study only had 50 participants, with somewhat skewed distributions of traits such as NYHA, age, and sex. These proportions were generally in line with what is typical of HF, but analyzing them as one group may have affected the quality of the results. There were 35 participants in NYHA II and 15 participants in NYHA III. Nearly 70% of the group was over 50 years of age, and only 2 participants were below 30 years of age. While 50 participants is a strong sample group for this kind of research, and many of the other studies in this area had similar groups, the analysis would have been stronger with a larger group. The main indication of this is the fact that the subgroup analysis of CPS results versus daily activity in elderly participants gave much stronger correlations. Lifestyle varies with age, especially before and after the age of retirement. Exercise habits of a participant in his or her twenties will differ from those of a middle-aged working adult, and from a retired senior.

The group also only included six women, or 12%. Only one female participant was in NYHA III. HF affects more men than women, and the clinic had very few eligible and interested female patients. Other studies also included fewer women than men, but still had a slightly higher
proportion of women than this study. Jehn et al\textsuperscript{47} had 12 out of 50, Dontje et al\textsuperscript{24} had 20 out of 68, and Izawa et al\textsuperscript{49} had approximately 30 out of 157. As percentages this is 24%, 29%, and 20% respectively, all higher than the proportion in this study. Differences associated with sex were corrected for by using percent of predicted results, although it would be preferable to rely on raw results instead of calculated ones.

6.3.2 Analysis limitations and assumptions

The analysis relies on the assumption that all the collected step data for each participant is directly relatable to the results from their CPS. While this is a reasonable and necessary assumption, some participants could have been having a good or bad day that may have offset their exercise test results. Similarly, events during the two-week study period may have caused some participants to exercise more or less than they would normally, such as meeting a deadline or being on vacation. Participants were also recruited throughout the year, in fall, winter, and spring. Quick analyses did not reveal any indication that start date strongly affected the results, but this would be difficult to show considering the many variables involved in comparing different participants. To understand this particular effect, it may be necessary to follow individual patients for one year or more, making sure to correct for any health-related or lifestyle changes.

The analysis presented here is somewhat specific to the TGH cardiac clinic. Some of the outputs from the CPS are fairly universal, but the tools and practices used may vary slightly between clinics. Some clinics may use treadmill-based exercise tests, which do not have watts as an output like a stationary bicycle. The use of predicted values may also differ. Other clinics may use NYHA classification differently. All participants were recruited from the same clinic, and an understanding of how different clinics operate is not included in the present work.

The experience from the study provided a general understanding of the usability of the consumer wristband activity tracker for HF monitoring as it would appear from the clinician’s end, but the study did not include a post-study questionnaire. A short follow-up questionnaire could have been sent to participants after the study ended to measure key aspects of the usability of the device from the perspective of the participants. This would have helped identify whether things like the daily charging, remembering to wear the device, and using the clasp on the device were
substantially disruptive to the patients’ lives. This was a missed opportunity to provide further evidence of the benefits of using a consumer device for this application.

A 6MWT could have been performed upon recruitment in addition to the CPS. Many patients performed a 6MWT in the clinic in between other tests and their appointment with their physician. Including the 6MWT would have made the recruitment process of the study slightly more complicated, but the 6WMT is a well-studied tool for evaluating functional capacity in HF, and many of the similar studies included the 6MWT in their analysis. For example Jehn et al performed baseline and final 6MWTs in the clinic with monthly tests performed independently by the patient at home. The device used in that study has a button to start a 6MWT and stops recording after 6 minutes. As part of this study we could have also asked participants to perform an unsupervised 6-minute walk at home a few times during the study. Participants could have been given a form to record when they did the 6-minute walk. This would have violated the concept of reducing requirements on patient, but it is not a big imposition and may have provided further evidence to support the use of consumer activity monitors for remote HF monitoring. In a practical application it is conceivable that patients could be expected to perform a periodic standard exercise like a 6-minute walk as a reference measurement.

6.3.3 Device limitations

While the concepts learned from this study apply to any activity tracker, the results themselves are restricted to the technology used. The algorithms used to estimate steps from accelerometer data captured by the wearable device are proprietary and kept private. At their core, all devices function by counting steps based on repetitive movements of sufficient magnitude, but sensitivities may vary between hardware, and special algorithms may be used to produce more accurate results. Different devices have been shown to produce different step counts in several investigations. Of course, knowing the exact number of steps is not the goal, but rather to estimate and compare activity levels between patients and more importantly, within a single patient’s activity history to understand how their condition is changing. Any activity tracker would provide an estimate of physical activity, but one cannot compare data from two different kinds of activity monitors running on different hardware or software without additional data to correct for the differences between the devices. Studies comparing activity trackers often compare manual step counts to the outputs from multiple activity trackers worn by an individual
on a treadmill. However, in real-world daily activity, the extra movements of the body from turning and interacting with objects in our environments affect step counts as well, especially for wristband devices. Different placements on the body give different results for step counts, so it is important to be consistent. It is worth noting that even if step tracking were perfectly accurate, small differences probably would not affect a clinical decision, since there are many extraneous variables involved.

As mentioned previous, one key limitation of the analysis is the inability to distinguish zero steps from not wearing the device. Participant adherence with regard to wearing the wristband could be influenced by HF severity. A participant who was not feeling well may be less likely to wear the wristband. However, personality and lifestyle are probably bigger factors when it comes to remembering to wear the wristband. For example, a very busy person may be more likely to forget their wristband in the morning. Other studies used devices that could track when they were being worn, with minimum requirements for daily wearing. Jehn et al\textsuperscript{47} required participants to wear the device for at least 12 hours per day for 8 consecutive days. Dontje et al required participants to wear the device for 48 hours without taking it off. While these requirements may have improved the quality of the final results, they impose limitations on patients that are not realistic if the tools were used in a practical setting for long term remote monitoring.

Future studies should try to use devices that can detect whether they are being worn. This should be very simple using an accelerometer sensor, as a completely stationary device sitting on a steady surface like a table would much less variability in an accelerometer than a person sitting in a chair or lying down. Other devices use sensors to track other physiological data like heart rate, and this could be repurposed to indicate whether the device was being worn. However, the daily adherence rates of participants in this study were still quite high. Considering the time of the first and last minute with non-zero steps for each of the over 700 collective days across all participants, the median difference in time is approximately 17 hours, with a 1\textsuperscript{st} quartile of over 13 hours. This means that on most days, the participant was probably wearing the device for most of the waking day. Of course, participants could and probably did remove the wristband between those times, but this still gives some indication of the wearing adherence. Also, the best performing metrics were actually ones that were not frequency-based, like 99\textsuperscript{th} percentile steps in 1 minute. These metrics automatically ignore days with less activity, only considering minutes
with non-zero steps. Minutes with zero steps are ambiguous, but minutes with one or more steps are not.

When a participant knows his or her activity levels are being tracked, they may perform more activity than they would have otherwise. This is an unavoidable fact when measuring exercise during any test. However, the devices used in this study only gave participants an approximate count of their steps through the dots that light up if the participant tapped on the face of the wristband. All participants were wearing the same device. Participants may have exercised more than they would have if they were not being monitored, but this would only have been beneficial to the participant and provided more data for the study.

The Fitbit Flex was also somewhat difficult to use, which may have affected how often participants wore it. The clasp is difficult to insert into the holes on the wristband, especially for older individuals with weaker or unsteady hands. During recruitment, many participants required assistance getting the wristband on. The insertion and removal of the tracker from the wristband and the charger was also unintuitive. Future studies should consider using devices that are slightly simpler, possibly without the requirement of removing parts to charge them or a need for a clasp.
7 Conclusions

Readmissions in HF are high and medical costs of HF are primarily from hospitalizations. Clinicians need new tools to better help patients, especially between visits when symptoms of decompensation that are ignored may lead to rehospitalizations. Improvements in telemonitoring have made it much easier for patients to stay connected with their medical support team. However, these systems have yet to include exercise monitoring, even though exercise capacity is a central part of HF management. Although there is plenty of research on activity monitoring, they rely on tools that do not fit well into patient lives or clinical workflows.

This project has shown that activity monitoring with a consumer wristband activity tracker is feasible and correlates with clinical exercise tests. Furthermore, this project has developed new metrics based on the format of the data outputted by consumer devices. These metrics do not rely as heavily on patient adherence as frequency-based metrics such as steps per day. These consumer devices are inexpensive, accessible, easy to use, reliable, and support patient self-monitoring.

7.1 Summary of Contributions

This thesis provides validation of the use of a consumer wristband activity tracker for remote monitoring of HF patients. Daily activity was related to standard quantitative clinical exercise tests and to more qualitative medical assessments of HF patients. This thesis explored many different combinations of exercise test and step data variables to find the best and most useful relationships. Daily activity is highly variable and uncontrolled, so one might not expect steps counted to relate to strictly controlled and supervised clinical exercise tests. However, this project has shown that activity monitoring has significant potential to supplement clinical exercise tests. Previous studies have shown this as well, but not using a consumer wristband activity tracker and usually not with a study design that simulates how these tools would be used in practice. The 99th percentile steps in 1 minute and the most steps in 10 minutes are simple to calculate and do not require long monitoring periods or the imposition of requirements for wearing the device. These peak activity metrics have not been proposed elsewhere; instead other studies have relied on measurements of daily total steps or walking time, which require consistent wearing of the tracking devices to be able to compare participants. By contrast, 99th
percentile steps in 1 minute and the most steps in 10 minutes only require a reasonable sample of a participant’s active day. Other step data outputs such as daily total steps, daily total walking time, and proportion of active time spent at a certain pace, also showed useful correlations with CPS data and NYHA classification.

This thesis also provided evidence of the usability and reliability of the Fitbit Flex activity tracker by simulating remote HF activity monitoring. This is also provides some indication that other similar consumer devices are sufficiently reliable to be used in this application. Although some technical issues were experienced, overall the devices performed well and were simple to use and collect data from. There is room for improvement, but new versions of these kinds of devices are frequently released that provide more features and better performance. The present work is a necessary step to moving forward with future studies that will apply and expand on these concepts.

### 7.2 Future work

Additional statistical analyses on the current data to further refine the results will be performed. This includes robust regression to identify outliers and influential observations. Additional variables, such as age, will be included with the NYHA classification logistic regression analysis to improve the quality of the prediction model. Additionally we will look at the daily variability in activity levels and peak step activity compared with the metrics that were based on the full study period. We can also try to determine the minimum amount of monitored active time that is necessary to achieve an accurate estimate of a patient’s exercise capacity.

Future studies will expand on the results from this project to look at how daily steps from these consumer devices changes over time, and how the data correlates with patient outcomes and other signals of patient status, such as weight and blood pressure. This project has established that consumer devices can measure daily physical activity in HF patients and this can be related in a variety of ways to exercise capacity. Future studies will draw ideas about methodology from this work so they can use activity trackers that can provide information on whether the device is being worn, and perhaps are easier to use. These studies will use metrics like 99th percentile steps in 1 minute or best 10-minute walk to analyze patient trends on a weekly basis over several months or longer. Prediction of NYHA classification from activity data could also be
incorporated using similar approaches.

Future studies should also consider the clinical aspect of remote activity monitoring; specifically how these techniques should be integrated into existing workflows. Activity data could be easily integrated into telemonitoring systems. Consumer devices come with all the data collection and transmission needs already built in, and is inexpensive and easy to acquire. The system would simply need to subscribe to and process the data from a patient’s activity tracker by authorizing the account paired with the tracker. The Centre for Global eHealth Innovation at the UHN is developing a telemonitoring platform called Medly, which has a mobile application that collects weight and blood pressure via Bluetooth-enabled devices and asks patients questions about how they are feeling. Activity tracking would be a useful addition to that system. Medly has a secure dashboard so clinical staff can review patient performance and receive alerts if patients report symptoms that require attention. How clinicians would incorporate daily activity data into their decision-making needs to be properly understood before these approaches can be put into practice, but this thesis has shown that step data correlates to clinical measurements that are already well understood with respect to evaluating HF. This is just the first step in what will eventually result in the inclusion of daily activity monitoring in practical telemonitoring.

7.3 Long term goals
Ultimately, step data metrics can be integrated into clinical telemonitoring and provide continuous information about a patient’s activity that can be related to their exercise capacity and how they are feeling. These tools can help patients become more aware of their activity by providing objective quantification, and may help them meet prescribed levels of activity. In between visits, clinicians have limited to no information about how a patient is feeling physically. Telemonitoring systems that allow patients to enter in symptoms and measure weight and blood pressure are very useful tools, but physical activity is missing from the telemonitoring toolkit. Limitations of physical activity is a strong indicator of HF severity, and being able to detect subtle changes in physical activity will enable quicker reaction to changing symptoms, or give clinicians more information when patients come in for appointments. Clinicians will have access to actual data to help them understand how patients are doing when they are at home, reducing reliance on subjective means of evaluating activity.
Consumer activity trackers also have significant potential throughout medicine, as physical activity is an important component of a wide range of conditions. These new tools will eventually form a central part of not only monitoring and management of diseases, but prevention as well. We must continue to explore and validate these tools, designing approaches for their use and helping to drive research into clinical practice to begin improving the lives of patients everywhere and helping clinicians be as effective as possible.
8 References


12. Types of Heart Failure. *American Heart Association* (2015). at <http://www.heart.org/HEARTORG/Conditions/HeartFailure/AboutHeartFailure/Types-of-Heart-Failure_UCM_306323_Article.jsp>


15. About Heart Failure. *American Heart Association* (2015). at <http://www.heart.org/HEARTORG/Conditions/HeartFailure/AboutHeartFailure/AboutHeartFailure_UCM_002044_Article.jsp>


42. Vooijs, M. *et al.* Validity and usability of low-cost accelerometers for internet-based self-


51. Wasserman, K. *Principles of Exercise Testing and Interpretation: Including Pathophysiology and Clinical Applications.* (Lippincott Williams & Wilkins, 2005).

Appendix A  CPS Summary Report Example

Toronto General Hospital
Heart Function Clinic
Cardiopulmonary Stress Test
Summary Sheet

Date
Age
Height
Weight

REST DATA
SBP  110
DBP  70
HR   52
O2 sat 100
FEV1 2.65
% Pred  74
FVC  3.4
% Pred  73

EXERCISE DATA
SBP  140
DBP  76
Heart rate 133
1 min rec HR 96
Drop in 1 min 37
O2 sat 100
Time 14:03
Watts 140
%predicted 90
Termination
Symptoms Y/N
1= leg fatigue
2= dyspnea
3= chest discomfort
4= HR/BP response
5= THR
6= presyncope
7= ECG changes
8= poor motivation

CP DATA
VO2 (ml/kg/min) 22.7
Predicted 31.4
% Predicted 72
VO2 (L/min) 1.534
Predicted 2.12
% Predicted 72
Peak VE 63.1
Peak VO2 (L/min) 1.94
VE/CO2 Peak 33
VE/CO2 @ AT 28
AT (ml/kg/min) 16
% meas. peak VO2 .71
% pred. peak VO2 50
Peak RER 1.26

Rest ECG
Rhythm: Sinus
A.Fib/flut
APB’s
VPb’s
Paced
Biventricular

Q-waves: anterior
inferior
lateral
posterior
Bundle Branch
LBBB
RBBB

LVH voltage: no
yes
ST/T abnormalities no
yes
leads:
Appendix B  Study consent form

CONSENT TO PARTICIPATE IN A RESEARCH STUDY

Study Title: Validation of a wearable activity tracker for the estimation of heart failure severity

Investigator: Joseph Cafazzo, PhD, P.Eng
Lead, Centre for Global eHealth Innovation, UHN
Telephone: 416-340-4800 ext. 3634

Co-Investigator: Dr. Heather J Ross, BSc, FRCPC, MD, MHSc
Affiliate Scientist Division of Experimental Therapeutics - Cardiovascular Toronto General Research Institute (TGRI)

Study Coordinator: Simon Bromberg
Centre for Global eHealth Innovation
Telephone: 416-340-8926
simon.bromberg@uhnres.utoronto.ca

Please note that the security of e-mail messages is not guaranteed. Messages may be forged, forwarded, kept indefinitely, or seen by others using the internet. Do not use e-mail to discuss information you think is sensitive. Do not use e-mail in an emergency since e-mail may be delayed.

FUNDING
This is a National Science and Engineering Research Council (NSERC) partnered project.

INTRODUCTION:
You are being asked to take part in a research study. Please read this explanation about the study and its risks and benefits before you decide if you would like to take
part. You should take as much time as you need to make your decision. You should ask
the study doctor or study staff to explain anything that you do not understand and make
sure that all of your questions have been answered before signing this consent form.
Before you make your decision, feel free to talk about this study with anyone you wish
including your friends, family, and family doctor. Participation in this study is voluntary.

BACKGROUND:
Doctors test a patient’s ability to perform exercise to measure how bad the patient’s
heart failure is, and how well he or she is recovering. A heart failure patient who is not
doing well is probably going to be taking fewer steps throughout the day, because the
heart has trouble responding to exercise loads. Lab tests like the one you are doing
today are an excellent way to test your heart health, but unfortunately this is not a
typical form of exercise and can only be done once in a while when a patient comes in
for a test. Wearable activity trackers can track physical activity all-day and may allow
better monitoring of heart failure outside of the hospital. This study will use a tracker
worn on the wrist like a watch to count steps.

A wristband activity tracker is an electronic device worn on the wrist that uses electronic
sensors to track physical activity, primarily by counting the number of steps taken
throughout the day. This is similar to a pedometer or step counter, except that it
provides a more detailed picture of daily activity. Note, the device is not able to track
where you go or what you do, only how many steps you take. The activity data collected
by the wristband is sent wirelessly to a mobile phone for analysis.

PURPOSE:
The purpose of this study is to relate physical activity as recorded by a wristband tracker
to heart failure. This study may lead to a tool that will enable doctors to keep track of
their patients’ daily activity remotely, to help with treatment planning and to help alert
the doctor if anything is wrong. Currently there is no easy way for doctors to measure and track patient activity outside
of the hospital. This study will show that activity trackers can be used to monitor heart
failure, and will be an important step towards providing clinicians with up-to-date and
accessible information about their patient’s recovery.

You are being asked to participate in this study because you have been previously
diagnosed with heart failure.

The monitoring technology that will be used in this study is not intended to modify or
encourage you to modify your activity levels in any way. This study should not affect
your treatment in any way. Nevertheless, monitoring heart failure through wristband
activity trackers is experimental, meaning that it is not routinely used in patients’ care.
Up to 100 people will participate in this study at UHN and it will take less than 1 year to complete.

**STUDY DESIGN:**
All participants will be given a wristband activity tracker, mobile phone with a data plan, charging cables, a log sheet, and a box for mailing the materials back at the conclusion of the study period. The phone will be used to wirelessly extract activity data from the wristband and automatically and anonymously upload that data to a secure online portal. The phone will be locked and participants are requested not to try to use the phone for any reason. The participant should plug the phone in a safe place in his or her home that they access daily and that has a strong cellular signal and leave it there for the duration of the study. The participant will wear the wristband on their wrist for 2 weeks (removing it for short periods when necessary, such as bathing) and then return it along with the other materials using the packaging provided.

Data from a wide range of patients with heart failure will be used along with the data of other participants to relate heart failure to activity measured by a wristband tracker.

**PROCEDURES:**
You will be asked a few questions about you and your everyday physical activities. Then you will be shown how to use the wristband and phone. You will be expected to wear the wristband every day during the study, removing it for short periods of time as needed. The tracker will need to be recharged at least every 5 days.

You will be in this research study for 2 weeks.

The wristband, mobile phone, and cables will not be available after the study is complete.

**RISKS:**
No major risks are associated with wearing the wristband or the mobile phone. If you experience any discomfort or skin irritation due to the wristband, please discontinue wearing the wristband and contact the study coordinator.

Activity data will never be linked to your personal information.

Some potentially sensitive information will be collected in the initial questionnaire.

Potential inconvenience may include wearing and caring for the wristband (such as remembering to take it off before a shower and putting it back on afterwards), as well as ensuring it is mailed back in a timely fashion. The wristband is comfortable and should be similar to wearing a wristwatch.
Please call the study coordinator if you have any side effects even if you do not think it has anything to do with this study.

**BENEFITS:**
You will not receive direct benefit from being in this study. Information learned from this study may help improve remote monitoring of heart failure patients in the future.

**REMINDERS AND RESPONSIBILITIES:**
It is important to remember the following things during this study:

- Ask the research team about anything that worries you.
- Tell the research team if you change your mind about being in this study.
- Care for the study materials properly; Do not subject them to extreme temperatures or pressures, twist, bend, or cut any part of the wristband. Please avoid dropping the materials. Do avoid getting the mobile phone and cables wet.
- Only you may wear the wristband during the study period. The wristband and other materials must remain in your possession during the study. Your friends and/or family are only permitted to interact with the provided materials for the purpose of helping you use them.
- Please leave the phone plugged in at all times in a safe part of your home that you access on a daily basis and where it can receive a good cellular signal.
- The phone that you have been provided with is intentionally locked, and you will not need to use it directly at any point during the study.
- You must return the device and the log as soon as possible on the final date of the study period. You will receive a call to remind you to return the device.
- The study coordinator may contact you during the study if he notices a problem or has any questions. You may be asked to extend or shorten your study period depending on the situation.

**Alternatives to being in the study**
You do not have to join this study to receive treatment for heart failure. This study should not impact your care. Consult with your physician for the usual standard of care.

**CONFIDENTIALITY:**

**Personal Health Information**
If you agree to join this study, the study team will ask you for your personal information relevant to the study. Personal information is any information that could identify you and includes your:

- Name
- Phone number
- Mailing address
- Year of birth
- Sex
- Weight
- Height
- Medical history related to your heart failure diagnosis
- Normal activity levels

Your participation in this study will also be recorded in your medical record at this hospital. This is for clinical safety purposes.

Research Information in Shared Clinical Records
If you participate in this study, information about you from this research project may be stored in your hospital file and in the UHN computer system. The UHN shares the patient information stored on its computers with other hospitals and health care providers in Ontario so they can access the information if it is needed for your clinical care. The study team can tell you what information about you will be stored electronically and may be shared outside of the UHN. If you have any concerns about this, or have any questions, please contact the UHN Privacy Office at 416-340-4800, x6937 (or by email at privacy@uhn.ca).

The following people may come to the hospital to look at the study records and at your personal health information to check that the information collected for the study is correct and to make sure the study is following proper laws and guidelines:

- Representatives of the University Health Network (UHN) including the UHN Research Ethics Board

The study team will keep any personal health information about you in a secure and confidential location for 10 years. Only members of the research team will have access to the data.

Study Information that Does Not Identify You
Any information about you will have a code and will not show your name or address, or any other information that directly identifies you.

All information collected during this study, including your personal health information, will be kept confidential and will not be shared with anyone outside the study unless required by law. You will not be named in any reports, publications, or presentations that may come from this study.

**VOLUNTARY PARTICIPATION:**
Your participation in this study is voluntary. You may decide not to be in this study, or to be in the study now, and then change your mind later. You may leave the study at any time without affecting your care. We will give you new information that is learned during
the study that might affect your decision to stay in the study.

You may refuse to answer any question you do not want to answer, or not answer an interview question by saying “pass”.

**WITHDRAWAL FROM STUDY:**
The Researchers will not ask you to return the study materials before the end of the study period for any reason.

If at any time during your study period you decide you do not want to participate anymore, please contact the study coordinator with the contact information included with your instructions and return all materials using the return package provided.

If you leave the study, the information that was collected before you left the study may still be used in order to help answer the research question. No new information will be collected without your permission.

**Costs and Reimbursement:**
You will not have to pay for any of the materials used in this study, nor will you have to pay for the postage to return them. You will not incur any expenses as a result of your participation in this study.

Upon successful completion of the study and return of the study materials you will receive a cheque via mail for $50 as a thank you for your participation.

**RIGHTS AS A PARTICIPANT:**
If you are harmed as a direct result of taking part in this study, all necessary medical treatment will be made available to you at no cost.

By signing this form you do not give up any of your legal rights against the investigators or involved institutions for compensation, nor does this form relieve the investigators or involved institutions of their legal and professional responsibilities.

**CONFLICT OF INTEREST:**
The research team has no conflict of interest to report.

**COMMERCIALIZATION:**
The research team and/or others intend to claim sole ownership of any research results consistent with this consent. You will not receive any financial benefit that might come from the research.

**QUESTIONS ABOUT THE STUDY:**
If you have any questions about the study please email our study coordinator, Simon Bromberg at simon.bromberg@unhres.utoronto.ca.
If you have any questions about your rights as a research participant or have concerns about this study, call the Chair of the University Health Network Research Ethics Board (UHN REB) or the Research Ethics office number at 416-581-7849. The REB is a group of people who oversee the ethical conduct of research studies. The UHN REB is not part of the study team. Everything that you discuss will be kept confidential.

You will be given a signed copy of this consent form.

**CONSENT:**
This study has been explained to me and any questions I had have been answered. I know that I may leave the study at any time. I agree to take part in this study.

___________________________  ___________________  __________________
Print Study Participant's Name  Signature  Date

My signature means that I have explained the study to the participant named above. I have answered all questions.

___________________________  ___________________  __________________
Print Name of Person  Signature  Date

Obtaining Consent
Appendix C  Study interview form

This form is to be used as a guideline for the recruitment interview and is to be filled out by the recruiter. This form is not intended to be given to the participant like a survey.

Study ID:   __________  Date:  ____________________

Age:  __________  Weight:  __________  Height:  __________

Sex (circle one):  Male  Female

Handedness (circle one):  Right-handed  Left-handed

Wristband preference:  Right Wrist  Left Wrist

Date of original HF diagnosis:  ____________________

Heart Failure Treatment to date:  (surgery, medication, diet / lifestyle changes)

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Most recent NYHA (New York Heart Association) Classification (circle one):

    I    II    III    IV

Regular physical activities undertaken (e.g. daily walks, swimming, bicycling)

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Contact Information (separate page)

Date: ____________________________

Recruiter: ________________________

Participant

First Name: ________________________

Last Name: ________________________

Phone number: ____________________

Mailing Address: ____________________

________________________________

________________________________

________________________________

Email: ____________________________

Note: Phone number or regularly checked email needed to contact participant about syncing issues etc. Will not be used unless necessary.

Mailing address needed to send participant gift card
Appendix D Study instructions

Thank you very much for agreeing to participate in this study. Your contribution will be an important part of this project.

Once you have gone over the study instructions, signed the consent forms, and completed the introductory questionnaire you will be provided with:

- 1 Fitbit Flex wristband and activity tracker
- 1 Fitbit charging cable
- 1 Android mobile phone
- 1 micro USB Android phone charging cable
- 1 USB wall charging adapter
- 1 log table
- 1 prepaid return package for mailing via FedEx

Fitbit Flex
The Fitbit Flex is an electronic device that is worn on the wrist and tracks steps throughout the day, like a pedometer. Wearing it is as safe as wearing a wristwatch and should not affect you in any way. The Fitbit wristband wirelessly communicates via Bluetooth with the mobile phone, which extracts step data and uploads it to a portal so we can access it. This data will be completely anonymous.

The tracker will need to be charged at least every 5 days for up to 3 hours. To avoid missing activity data, please charge overnight every 3 days. Charging instructions are attached.

Mobile Phone
The Moto G X is an Android smartphone that connects to the Internet using cellular data. You will not be able to use the phone for anything other than collecting data from the wristband. Find a place in your home where you will go frequently, such as in your bedroom, where it can be plugged in and left safely throughout the study. Do not turn off the phone at any point during the study. Line of sight is not necessary for syncing. While you do need to wear the wristband throughout the day, the phone can be left at home, as long as you are within 5 meters of it at least once per day. Make sure that the phone is able to get a cellular signal. You can check whether the signal and that the battery is fully charged by pressing the power button on the side briefly and looking for the symbols showing in the figure below at the top right of the phone’s screen.

If the signal strength indicator (left side) has only one or two bars, you may need to find
another spot to leave the phone.

The electronic devices are designed to withstand normal conditions, but it is important that you be careful not to damage it as you would with a wristwatch. In addition, avoid subjecting it to extreme heat or cold for extended periods of time outside the range of -20° to 45° C.

**Wearing the Wristband**

Wear the wristband loosely enough to allow air circulation. Clean and dry the wristband regularly, particularly under the band. Please do not remove the tracker from the wristband enclosure except for charging.

Please contact the study coordinator and discontinue wearing the wristband if the material feels uncomfortable or causes any irritation on your skin.

You, and only you, will wear the band for the duration of the study on the same wrist. You may remove it for short periods of time. If you must spend over 1 hour without wearing the band, please make a note of that in the log provided. If you are engaged in physical activity significantly outside of your norm, or extended periods of inactivity for any reason, please also record that in the log.

**Important:** You should only perform physical activity within your capabilities and within what you and your physician have agreed is safe for you. We are interested in studying your normal activity levels, and in no way are we encouraging you to modify your current lifestyle. All of your activity data will be kept anonymous.

Note: the band is only able to track steps; it cannot identify where you are or what you are doing.

Showering or swimming while wearing the wristband is not recommended, but the wristband is water-resistant. You can wear the band while you sleep.

Please note that the band will be thoroughly cleaned before and after you use it. If you need to clean it during the study, rinse it with warm water and mild detergent.

**Upon Completion of the Study**

Prior to the final day of your study, you should receive a phone call or e-mail from the study coordinator. The study coordinator will remind you to return the study materials via FedEx pickup or dropoff at a time convenient for you. Please try to return the materials within a day or two of the completion of your study.

Call FedEx to arrange a pickup time after the final day of the study period that is
convenient for you. Call FedEx at 1.800.GoFedEx (1.800.463.3339), and say “speak to a representative” when prompted or dial 0. Tell them you would like to arrange a pickup and choose a time and location that is convenient for you. If you prefer, you can also drop off the package at an authorized FedEx location, or arrange with the study coordinator to return the package at Toronto General Hospital. If you encounter any problems, please contact the study coordinator.

Remove the wristband, record the date and time on the corresponding spot at the top of the log, seal all of study materials (including all of the materials listed on page 1 of these instructions) in the return package and give it to the FedEx representative when he or she arrives. Please wrap the materials carefully in the provided packaging materials, especially the phone and Fitbit tracker.

If you have any questions or concerns please do not hesitate to call or email the study coordinator.
FAQs

1. Can I swim / shower while wearing the wristband?
   Yes, but if you prefer to take it off, then do so. Do not wear the wristband while diving. Keep in mind that the wristband may need to be dried. The phone and chargers are not waterproof.

2. Can I switch the wristband to my other wrist?
   Please try to be consistent with which hand you wear the wristband on, preferably your non-dominant hand.

3. What do I do if there’s something wrong with the wristband / the phone?
   Please contact the study coordinator.

4. How will I know if it’s tracking / monitoring properly?
   We will try and contact you if we think there is a problem with the tracking. To check whether the Fitbit is working properly, tap twice quickly and firmly on the face of the wristband. If it is working properly, it should show a brief animation of dots followed by 5 dots.

5. Why did my Fitbit Flex wristband vibrate? OR Why is the screen showing 2 blinking dots?
   You may have accidentally activated a mode that is not necessary for the study. To exit this mode, tap rapidly on the face of the wristband for 3 seconds. It should vibrate, show an animation of the dots across the screen, then show 5 dots, and finally go blank again.
Instructions for charging activity tracker

To charge your Flex, do the following:

1. Unclasp the wristband from your wrist and remove the tracker from the wristband.

2. Insert your tracker into the charging cable. Line up the tracker with the charging cable, putting the rounded end of the tracker into the charging cable at a slight angle. Use your thumb to push the tracker in and down into the charging cable until you hear a click.

3. Insert the charging cable into the provided wall charger and plug that into a wall outlet.

4. Your Flex's LED indicator lights will pulse to show the battery level every few seconds. Each solid indicator light represents progress towards the total charge. It may take up to 3 hours to fully charge your Flex. **When it’s charged, all 5 indicator lights will blink.**

5. Unplug the wall charger from the wall, and unplug the charging cable from the wall charger. Then gently pop the tracker out of the charging cable by pressing upwards on the exposed backing of the charging cable.
6. Insert the tracker back into your wristband and put the wristband back on the wrist you have chosen to wear it on.