SF 424 R&R and PHS-398 Specific
Table of Contents

PI: STRATH, SCOTT J	Title: Calibrating free-living physical activity characteristics across functionally-limited populations using machine-learned accelerometer approaches	
	FOA: PA16-167	
	FOA Title: DIET AND PHYSICAL ACTIVITY ASSESSMENT METHODOLOGY (R01)	
	Organization: UNIVERSITY OF WISCONSIN MILWAUKEE	
Senior/Key Personnel:	*Organization:*	*Role Category:*
Scott Strath Ph.D	University of Wisconsin Milwaukee	PD/PI

RESEARCH & RELATED Other Project Information

1. Are Human Subjects Involved? ☒ Yes ☐ No

1.a If YES to Human Subjects

Is the Project Exempt from Federal regulations? ☐ Yes ☒ No

If NO, is the IRB review pending? ☒ Yes ☐ No

2. Are Vertebrate Animals Used? ☐ Yes ☒ No

3. Is proprietary/privileged information included in the application? ☐ Yes ☒ No

4a. Does this project have an actual or potential impact on the environment? ☐ Yes ☒ No

5. Is the research performance site designated, or eligible to be designated, as a historic place? ☐ Yes ☒ No

6. Does this project involve activities outside of the United States or partnerships with international collaborators? ☐ Yes ☒ No

PROJECT SUMMARY/ABSTRACT

One in 5 U.S. adults are thought to be living with a disability/impairment, a complex and multifaceted condition affecting movement patterns with related medical care costs exceeding $300 billion annually. Precise and accurate assessment of physical activity (PA) and sedentary behavior (SB) in individuals with disability/impairment is essential to accurately measure PA/SB prevalence rates and effectiveness of behavioral based PA/SB interventions, and to fully elucidate PA/SB dose-response health relationships. Scientific progress has been made in this area with advanced analytics and data processing techniques applied to wearable accelerometers from laboratory calibration studies. There is a scientific need to extend calibration studies from fixed-duration laboratory simulated activities of daily living to free-living calibrations with natural observation and accelerometer algorithm training and validation. The aims of this proposal fill this essential scientific knowledge gap. The specific aims are: 1) To evaluate and refine machine-learned algorithms to predict energy cost and activity type during a 24-hr respiratory calorimeter stay; 2) To validate machine-learned accelerometer algorithms with field-derived, video-recorded direct observation; and 3) To validate machine-learned algorithms using the doubly labeled water technique. Our highly qualified research team will address the above aims by using brief translatable functional tests to cluster movement-impaired populations into groups of healthy, upper-body impairment, lower-body impairment, and upper- and lower-body impairment. Best practice free-living calibration protocols will then be used to train, refine, and evaluate functional clustered-specific accelerometer algorithms for predicting activity energy cost, activity type, activity transitions, and activity domain. The results of these proposed studies will for the first time provide an innovative and translatable approach to categorize and assess free-living PA/SB in persons with disability and movement impairment.

NARRATIVE

The accurate assessment of physical activity (PA) and sedentary behavior (SB) in populations with disability/functional impairment is essential for PA/SB surveillance, for determining effectiveness of behaviorally based PA/SB interventions, and for further elucidating dose-response relationships between PA/SB and health. The overall aim of this application is to extensively validate wearable accelerometers using advanced analytics and data processing in free-living naturalistic settings across categories of functional movement impairment, thus providing a convenient and translatable method to assess PA in diverse populations.

RESOURCES

UNIVERSITY OF WISCONSIN-MILWAUKEE

Description of the Institutional Environment:
The University of Wisconsin-Milwaukee (UWM) is the second largest University in the Wisconsin system and one of two Ph.D. granting research universities in the system. Located in an urban setting, it has an enrollment of over 28,000 students representing 48 states and 80 nations. It offers 186 degree programs, of which 93 are graduate programs, and has a faculty and instructional staff size of 1703. The University has several large-scale research institutes and numerous research libraries, laboratories, centers, and other units to form a well-connected network of research resources. At UWM, research expenditures have increased from $23 million in 2000-2001 to over $59 million in FY2014. Growing research expenditures has been a focal point of UWM during the past decade, along with dedicating campus resources to support research productivity. The University has extensive computer support which is all networked providing all faculty and students with resources that are easy to share. The University library system is interconnected with all surrounding institutes, and currently owns over 5 million volumes or works. There is ample support for researchers to gain full access to on-line journals and databases, all facilitating the research process. All faculty and staff have access to basic and extensive computer packages. Support for computing is provided by a dedicated staff with consultants responding to questions and concerns within a timely manner.

Notably, in January 2015, UWM was officially selected as a *"Community Engagement Classified Campus"* by the Carnegie Foundation for the Advancement of Teaching, and in February 2016, was classified as an R1 institution in the Carnegie Classification of Institutions of Higher Education™

The UWM, Department of Kinesiology is one of the most recognized departments on campus housed within the College of Health Sciences. The Department hosts 6 independent research laboratories, all with nationally recognized research foci. The Department has 15 fulltime faculty members, dedicated research coordinators, and 14 doctoral students.

The UWM-Center for Aging and Translational Research (CATR), will serve as the primary project administration site. Multidisciplinary in nature, CATR is comprised of faculty and staff from the health and social sciences with a research focus on the gerontological population. CATR provides access to two specialized support teams to facilitate optimal support for faculty. The administrative team provides institutional support for grants management (pre and post award), and is composed of a business manager, an administrative assistant, and student workers. The research core, which is composed of a research director, a database design specialist, a statistician, and a methodologist, provides grant submission assistance, database and web design services for data collection and management, and data analysis services. Thus, the CATR scientific environment promotes successful completion of projects by providing individualized full administrative support and access to a comprehensive set of research and data management resources.

Facilities:

Laboratory: The Physical Activity and Health Research Laboratory, affiliated with CATR, is located in the Department of Kinesiology, is over 2000 square feet in size, with an additional 600 square feet of adjoining office space to house all personnel necessary to carry out this proposal. The Laboratory comprises of a comprehensive body composition assessment core

and a physical activity assessment core. The Laboratory is also equipped to perform basic cardiovascular and metabolic assessments highlighted in this application. Exercise testing equipment includes 2 TMX425C TrackMaster treadmills, 3 Monark mechanically braked cycle ergometers, 1 Quinton electronic braked cycle ergometer, 1 Monark 881e arm crank ergometer, 2 ParvoMedics TrueOne Metabolic Measurement Systems, 1 Welch Allyn resting and exercising electrocardiogram system. The Body Composition Core Laboratory housed within the Physical Activity and Health Research Laboratory includes 1 Bod Pod Body Composition System (Life Measurement, Inc.), 1 Hydrostatic Weighing Tank with Residual Volume Measurement System (Exertech), 1 Dual Energy X-ray Absorptiometry Body Densitometer (GE Lunar Prodigy), a Bioelectrical Impedance Analyzer (Tanita TBF-612) and sufficient equipment to carry out full anthropometric measurements. The Physical Activity Assessment Core Laboratory, also housed within the Physical Activity and Health Research Laboratory includes 150 Actigraph Gt3X+ accelerometers, 20 charging stations, and a variety of other wearable motion sensors including an IDEEA monitor, 5 portable still image cameras (Sensecam & Narrator Clip devices), pedometers, and heart rate and global positioning system devices. As part of the Physical Activity Assessment Core the laboratory also has a customizable direct observation platform (Noldus Technology) and 4 dedicated PCs for coding. The core has developed already precoded videos and an extensive training module for the use of rigorous direct observation protocols. The laboratory also owns a portable metabolic measurement system (Cosmed K4b^2, Italy) and a computerized test bank of physical activity questionnaires. All major equipment items are available to support the successful completion of this project.

The Physical Activity and Health Research Laboratory has staff trained in all aspects of phlebotomy, physical activity assessment, body composition and anthropometric assessment, nutritional assessment, and basic cardiovascular assessment.

The Physical Activity and Health Research Laboratory has several Dell PC computers and 3 Dell laptop computers all available for physical activity assessment, data and statistical analyses. Computer software including statistical software, database, and general office applications are also available in the Laboratory.

Laboratory facilities coupled with University, Department, and Center resoucres creates an extremely supportive environemnt to successfully carryout the aims as outlined in this research proposal.

***Animal**:* Not applicable

***Clinical**:* Not applicable

***Computer**:* CATR faculty, through the Center's relationship with the College of Health Sciences, also have access to various University owned cyberinfrastructure resources, including the UWM High Performance Computing (HPC) Service. The HPC service supports a research cluster called Avi. Avi specifications include:

- 142 compute nodes (1136 cores total).
- Each node is a Dell PowerEdge R410 rack-mount server with two quad-core 2.67 GHz Intel Xeon X5550 processors and 24GB of system memory.
- one LSF scheduling node, a Dell PowerEdge R710 server, with two quad-core 2.67 GHz Intel Xeon E5520 processors and 24 GB of system memory.
- one IO node, a Dell PowerEdge R710 server, with two quad-core 2.67 GHz Intel Xeon E5520 processors and 24 GB of system memory.

- 7 Dell PowerVault MD1000 3Gb/s SAS attached expansion units providing 80TB of RAID 60 and RAID 10 storage. This storage is available to all nodes via NFS.
- each node has both a Qlogic DDR InfiniBand (16Gb/s) and a gigabit ethernet network interface.
- All nodes run Redhat Linux 5.3.
- In addition to the HPC Service, CATR faculty have access to a variety of off campus computing resources, including Red Cloud, Cornell's on-demand research computing service available by subscription. It provides virtual servers managed by the subscriber, and a software service for running large-scale Matlab models utilizing the Matlab distributed computing server and nVidia GPUs. Computing resources are also available through the University via XSEDE (formerly TeraGrid), SeWHiP, Open Science Grid, and Amazon EC2. The computational resources available are more than adequate for the completion of the proposed research.

THE UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS (CU-AMC)

Description of the Institutional Environment:

The Health Sciences Center of the University of Colorado Denver (UCD) underwent a relocation to the new Anschutz Medical Campus (CU-AMC) that was completed in 2009. This campus encompasses the Schools of Medicine, Nursing, Dentistry, Pharmacy, Public Health and Graduate Studies, as well as the University of Colorado Hospital and the Children's Hospital. A new Veteran's Affairs Medical Center is scheduled to be completed on this campus at the end of 2016.

Division of Endocrinology, Metabolism, and Diabetes

Office: Dr. Melanson maintains office space (-150 sq ft) in the Academic Office Building 1 **(A01),** which is located in in close proximity to the clinical and metabolic laboratory facilities. This building is behind the University Hospital, where the calorimeter is located, and next to the Leprino Building, where the CTRC Core Lab and NORC Energy Balance Assessment Core are located. The PI has an individual 150 sq ft office and access to a FAX machine and photocopier. Dr. Melanson has support from 2 full-time administrators who provide assistance to faculty for fiscal management of awards, ordering of supplies, and secretarial needs. Other university resources that are available for use include a machine shop, an electronics shop, and other necessary support services. In addition, CU-AMC is a major research institution with a large and talented faculty. Advice, assistance, and collaboration are readily available from experts in a wide range of research areas.

Wet laboratory: Dr. Melanson has -1500 ft^2 of shared wet laboratory space on the $5t^h$ floor of the RC2 Building which includes two separate closed rooms for microscopy and tissue culture. The laboratory has shared use of major resources (e.g., ice machines, autoclaves, walk-in cold room, dark room, etc) and dedicated space for -80C freezers and other storage. Dr. Melanson's wet laboratory also houses and maintains the OA-ICOS instrument. Dr. Melanson also possesses a license for the commercially available Post Analysis Software (Los Gatos Research Inc, Mountain View CA, Version 2.2.0.12).

Computers: All research personnel have Windows-based computers that are connected to local departmental servers that will provide the means for data transfer, electronic communication, and Internet access. All of the offices and laboratories described above have wired and wireless connectivity to institutional networks and are supported by the CU-AMC Information Technologies department. The Pls and their colleagues store all research data on central servers that are managed by the CCTSI Translational Informatics Core.

Colorado Clinical Translational Science Institute (CCTSI)

The CCTSI is supported by an ongoing (and recently refunded) NIH CTSA grant. The resources within the CCTSI that will be utilized by the proposed clinical trial include the adult outpatient CTRC and core laboratory located in the Leprino Building, the inpatient research unit within the University of Colorado Hospital (UCH), the Colorado Biostatistics Consortium (CBC), and the Translational Informatics Core.

Outpatient facilities and infrastructure: The outpatient CTRC includes 11 exam rooms, 3 metabolic testing rooms, a 3-station endurance testing laboratory, a strength testing laboratory, an ultrasonography laboratory, a 6-station phlebotomy center, and a 5-station infusion center. The adult CTRC unit provides personnel support (nurses, medical technicians, bio-nutritionists, medical technologists) for clinical visits and core laboratory assistance.

Inpatient facilities and infrastructure: The inpatient CTRC includes 4 ICU rooms, 8 step-down or medical/surgical rooms, a metabolic chamber, a procedure room, and a sleep laboratory. The facility is staffed 24/7 by nurses and medical support staff. The metabolic chamber (room calorimeter) is operated by the Energy Balance Assessment Core of the NORC (below), under the direction of Dr. Melanson.

Core Laboratory: The CCTSI Core Laboratory and sample storage facility are located adjacent to the outpatient CTRC. The laboratory is accredited by the College of American Pathologists (CAP). It has a pneumatic tube transport system that automates the transfer of samples to and from the UCH. The laboratory performs a number of specialized endocrine, inflammatory marker, lipid, and oxidative stress assays for approved clinical research protocols of the CCTSI.

Colorado Biostatistics Consortium (CBC): The CBC is the academic home for biostatisticians who provide support to clinical investigators within the CCTSI network. This arrangement allows investigators to 'buy' part-time support of additional junior-level biostatisticians as needed.

Translational Informatics: The centralized database for this study will be maintained on servers in the Translational Informatics Core. The CCTSI supports REDCap (a secure, web-based application designed exclusively to support data capture for research studies) for data require manual entry. Some limitations imposed by REDCap have encouraged us to develop our own paperless/electronic data capture system.

Nutrition and Obesity Research Center (NORC)

The UCD was first awarded a Clinical Nutrition Research Unit (CNRU) in 1995. This CNRU was continuously funded through 2010 when it was successfully converted into a Nutrition Obesity Research Center (NORC, 5 P30 DK048520-16) in a competitive renewal process. The NORC supports 3 Core labs. The Metabolic Core Services Center performs a range of metabolic and

molecular assays and gene expression by RT-gPCR on samples from both human and animal studies and is led by Dr. Jed Friedman. The Energy Balance Assessment Core (EBAC) laboratory provides equipment and support for the assessment of resting metabolic rate and body composition. The following equipment is available to support the proposed studies: 2 indirect calorimeters (Parvo Medics TrueOne 2400, Salt Lake City, UT); as well as a Dual Energy X-ray Absorptiometer (DXA, Hologic Discovery W version Apex 4.0.1 Hologic Inc., Bedford, MA).

Whole-Room Calorimeter: Dr. Melanson is the Director of the Whole-Room Calorimeter Core Laboratory, which is a sub-core to the NORC Energy Balance Assessment Core and is located in the CTRC. The calorimeter room is 12 feet x 12 feet and contains a regular hospital bed, desk, toilet, telephone, flat screen TV with a DVD player, and computer with wireless internet access. There are curtains over the windows for privacy. The room is equipped with a closed-circuit camera that can be viewed in the control room and at the nursing station. This camera permits the nurses to continually monitor subjects when they are residing in the room. However, the control for the camera is located inside the room, and subjects can turn the camera off anytime for privacy, for example, when using the bathroom. The room has a leak-free port through which blood samples can be obtained by having the subject extend their arm through the port. The nursing staff on the CTRC provides 24-hour coverage and supervision when subjects are in the calorimeter. Calibration, monitoring of the instrumentation, and data downloading and analysis are regularly performed by members of the research team. Supervision during the exercise sessions is performed by members of the research team.

UNIVERSITY OF MASSACHUSETTS AMHERST

Computer: The Research Computing Facility provides network support, system administration, software and hardware support to approximately 50 UNIX workstations and over 100 PC's and MAC's operated by the faculty, staff and graduate students of the Department of Mathematics and Statistics. The RCF also supports and manages peripheral equipment such as a variety of high-speed printers and scanners. In addition, the RCF provides system software and hardware consultation to the many end users of the Department.

Office: The University of Massachusetts, the Department of Mathematics and Statistics, and the University Library provide office space and library privileges to faculty, staff, research assistants and consultants.

Computer Cluster: The University of Massachusetts is one of five universities involved with the Commonwealth of Massachusetts in the Massachusetts Green High Performance Computing Center (MGHPCC). This center provides state-of-the-art computing infrastructure and a shared 5000-core computing system available free of charge to researchers across the UMass campuses. Each researcher is also allocated a 50GB home directory, with additional space available by request.

Animal: Not applicable

Clinical: Not applicable

EQUIPMENT

UNIVERSITY OF WISCONSIN-MILWAUKEE

The following essential equipment items are located within the University of Wisconsin-Milwaukee Physical Activity and Health Research Laboratory:

- 150 Actigraph Gt3x+ Accelerometers (Actigraph LLC) and software
- 20 Charging hubs and USB cords for Actigraph Gt3x Accelerometers
- Nold$_u$sTM Technology customizable direct observation platform, training manuals/procedures
- Physical activity assessment devices (stimm image camera capture — Sensecam/Narrator clip, heart rate monitoring, GPS)
- 2 ParvoMedics TrueOne Metabolic Measurement Systems (one with RMR dilution hood)
- Dual Energy X-Ray Absorptiometer (GE Lunar Prodigy) for the assessment of body composition
- Sphygmomanometers and stethoscopes
- Statistical analysis software
- Tension fitted tape measures
- Portable physician's scale and stadiometer
- Microsoft Access database software

Additional Equipment located in the Physical Activity and Health Research Laboratory

- A portable metabolic measurement system (Cosmed K4b^2, Italy)
- 1 GE CASE 8000 resting and exercising electrocardiogram system
- 2 TMX425C TrackMaster treadmills
- 3 Monark mechanically braked cycle ergometers
- 1 Quinton electronic braked cycle ergometer
- 1 Bod Pod Body Composition System (Life Measurement, Inc.)
- 1 Hydrostatic Weighing Tank with Residual Volume Measurement System (Exertech)
- Nutritionist Pro Software for dietary analysis
- YSI Glucose/Lactate analyzer
- Fume hood
- Full wet lab space with bench top preparation areas

The following essential equipment items are located within the University of Wisconsin-Milwaukee Center for Aging and Translational Research:

- Teleform software package on dedicated computer with dedicated scanner for forms based data collection and instrument design
- Database development utilizing various software packages, including Teleform and SPSS
- Participatory Recruitment Databases
- Assistance with recruiting and scheduling participants

- Extensive statistical software packages, including SAS, SPSS, STATA, Lisrel, Winsteps, and others

PHS 398 Cover Page Supplement

OMB Number: 0925-0001

1. Project Director / Principal Investigator (PD/PI)

Prefix:

First Name: Scott

Middle Name:

Last Name: Strath

Suffix:

2. Human Subjects

Clinical Trial? ☒ No ☐ Yes

Agency-Defined Phase III Clinical Trial?* ☐ No ☐ Yes

5. Human Embryonic Stem Cells

* Does the proposed project involve human embryonic stem cells? ☒ No ☐ Yes

PHS 398 Research Plan

OMB Number: 0925-0001

Please attach applicable sections of the research plan, below.

1. Introduction to Application (for RESUBMISSION or REVISION only)	
2. Specific Aims	1250-2_Specific_Aims.pdf
3. Research Strategy*	1251-3_Research_Strategy.pdf
4. Progress Report Publication List	
Human Subjects Sections	
5. Protection of Human Subjects	1252-5_Protection of Human Subjects.pdf
6. Inclusion of Women and Minorities	1253-7_Inclusion_of_Women.pdf
7. Inclusion of Children	1254-8_Children.pdf
Other Research Plan Sections	
8. Vertebrate Animals	
9. Select Agent Research	
10. Multiple PD/PI Leadership Plan	1255-11_Multiple_PI_Plan.pdf
11. Consortium/Contractual Agreements	1256-11_Consortium_Contractual.pdf
12. Letters of Support	1257-13_Letters_of_Support.pdf
13. Resource Sharing Plan(s)	1258-14_ Resource_Sharing_Plan.pdf
Appendix (if applicable)	
14. Appendix	

SPECIFIC AIMS

Accelerometer-based wearable sensors have become the standard for measuring physical activity (PA) and sedentary behavior (SB) due to their unobtrusive size and vast data collection capabilities. The scientific community continues to develop a strong understanding of their use and analytic processing to assess PA/SB in healthy populations. There has been a slow evolution of calibration standards: this has progressed from *traditional* laboratory-generated, linear regression-derived and intensity-demarcated "cut-points" to using simulated activities of daily living (ADL) combined with advanced analytical machine learning methods to predict both intensity and activity type. However, the use of accelerometers and data processing in non-healthy populations represents a significant knowledge gap. While some groups have developed disease-specific or age-specific algorithms, this work is limited because it largely adopts calibration protocols that mimic *early traditional* approaches in healthy populations, i.e. linear-regression/receiver operator characteristic binary outcome "cut-points" derived from lab-measured fixed-time duration start-stop ambulatory activities.

The use of accelerometer methods to assess PA/SB is predicated on the technology being a biomechanical movement sensor, able to distinguish different movement behaviors. Therefore, the use of accelerometry to assess PA/SB is not dependent on disease or age per se, but rather on movement patterns. In our recently funded R21 **(R21HD080828),** we clustered a heterogeneous sample of adults without movement disorders and patients with movement disorders into functional movement pattern impairment groupings of: a) none, b) upper-body, c) lower-body, or d) combined upper- and lower-body functional impairment. We have successfully developed translatable function tests to discriminate functional cluster groups and explored advanced analytic accelerometer algorithms for each function group using fixed-time ADLs simulated in the laboratory (see preliminary studies). The overall objective of this RO1 application is to reinforce these movement cluster groupings employing translatable functional tests, and to further calibrate the developed accelerometer algorithms using best-practice free-living data collection methods and advanced analytics across multi-accelerometer site placement. *Our central hypothesis* is that compared to algorithms derived in healthy populations and applied to all populations, movement pattern-specific accelerometer algorithms will have greater accuracy in classifying activities and predicting energy expenditure in free-living scenarios, scientifically advancing our ability to assess PA/SB in diverse populations in naturalistic settings.

AIM 1: Evaluate and refine machine-learned algorithms to predict energy cost and discriminate activity type during a 24-hr respiratory calorimeter stay. Hypothesis: Analytical models that are derived on physical function clusters using machine learning methods will produce more accurate and precise estimates of PA intensity and type classification compared with healthy population derived accelerometer algorithms applied to diverse populations. Method: After functional group assignment using translatable function tests, we will evaluate and refine functional group cluster-specific high resolution hip, wrist, and ankle accelerometer data machine-learned algorithms using min-by-min energy cost and activity type as determined from a 24-hr room calorimeter. Deliverable: Upon completion of this Aim we will have refined multiple time and frequency domain inputs for machine learning algorithms to process high resolution accelerometer data obtained from varied non-fixed start-stop time exercises and ADLs across healthy and diverse physical function-impaired populations.

AIM 2: Validate machine-learned accelerometer algorithms with field-derived, video-recorded direct observation. Hypothesis: Physical function-clustered, specific machine-learned accelerometer algorithms will demonstrate more accurate and precise PA intensity and activity type classification during free-living compared with healthy population derived

accelerometer algorithms applied to diverse populations. Method: We will test functional group cluster-specific high resolution hip, wrist, ankle accelerometer algorithms for free-living PA intensity/type using random day video-recorded direct-observation. Deliverable: Upon completion of this Aim we will have validated our functional clustered-specific accelerometer algorithms in natural environments.

AIM 3: Validate machine-learned algorithms using the doubly labeled water technique. Hypothesis: Physical function-clustered specific machine-learned accelerometer algorithms will produce more accurate and precise estimates of PA energy expenditure (PAEE) during free living compared with healthy population derived accelerometer algorithms applied to diverse populations. Method: We will test functional group cluster-specific high resolution, hip, wrist, ankle accelerometer algorithms for estimating PAEE using doubly labeled water method. Deliverable: Upon completion of this Aim we will have demonstrated the validity of our machine-learned functional clustered-specific accelerometer algorithms in natural environments to estimate PAEE.

Impact. Our results will provide much-needed clarity in the free-living approach to accelerometer calibration in heterogeneous populations. Results will provide high-impact valid estimates of PA/SB characteristics across diverse healthy, diseased, and movement-limited individuals. This will more clearly elucidate PA/SB prevalence rates, effectiveness of behavior-based interventions, and dose-response relationships to inform evidence-based guidelines and practice to prevent and/or manage disease conditions in diverse populations.

2. RESEARCH STRATEGY.

2.a. SIGNIFICANCE.

SCIENTIFIC PREMISE

2.a.1. Disability and impairment. As defined by the World Health Organization, disability is a broad umbrella term that covers impairment, activity limitation, and participation restriction.' Disability therefore represents a multifaceted interaction between a person and the environment/society in which they live.[2, 3] A physical impairment can involve an individual who moves with a deviation from normal, such as not being able to make a muscle or limb move efficiently, and this deviation limits or alters activities of daily living (ADL) or instrumental activities of daily living (IADL). A vast array of different diseases or conditions therefore fall under the term of disability, or physical impairment. For instance, cancer survivors report a 3-fold higher level of disability relative to the general population.[4] By specific age group, cancer survivors are more likely to report limitations in ADLs (personal care needs, household activities) and functional limitations (i.e., walking and social activity participation). Results of large national surveys also identify cancer as being most associated with disability and impairment.[5] Other diseases also constitute linkages with disability including, but not limited to, arthritis, multiple sclerosis, Parkinson's disease, stroke, cerebral palsy, and low-back pain. Diseases or conditions could be acute or chronic in nature, so disability classification could be transient - in and out of classifications of impairment, or they could see progression of impairment over time. National U.S. statistics report approximately one in five adults live with a form of disability.[6] The prevalence of reported disability increases with age, with medical care disability-related costs consistently exceeding $300 billion annually in the U.S. for the last two decades.' The number of individuals reporting a disability will likely increase, along with medical costs and public health service need, as more and more individuals enter old age.

2.a.2. Physical activity and sedentary behavior: Importance to health. Regular engagement in physical activity (PA) and reduced levels of sedentary behavior (SB) have been clearly associated with lower rates of all-cause mortality, coronary heart disease, type 2 diabetes, and other characteristics of metabolic and functional health.[5, 9] Our understanding of the benefits of PA and reduced SB on health is not solely restricted to individuals *without* physical impairment; Indeed, evidence supports the essential need for PA in individuals with disability, guided by some, albeit limited, epidemiologic and interventional evidence.[10-27] Even though there is a high prevalence of individuals with disability/impairment, and studies report on the benefits of PA to health for those with disabling or impairing conditions, research into precise and accurate methods to assess PA and SB in disabled/impaired populations represents a significant scientific knowledge gap. Advancement in the science of PA/SB assessment specific to populations with disability/impairment is essential to further elucidate dose-response relationships, to accurately estimate prevalence estimates of health-enhancing PA and SB reductions, and to identify interventional and therapeutic efficacy and change over time.

STRENGTHS/WEAKNESSES OF CITED STUDIES.

2.a.3. Physical activity and sedentary behavior assessment in disabled/impaired individuals. Self-report PA and SB surveys are available specific to population,[25, 29] type of setting,[3o-34] age,[35-37] and disability status.[35] Surveys have been improved upon over the years but are still plagued by memory recall error and bias, limiting accuracy and precision.[33, 39-41] Objectively assessing PA and SB through wearable devices has predominated over the last 10-20 years, due to their unobtrusiveness and rapidly expanding technology capabilities.

Technological advancement in wearable accelerometer monitoring now permits researchers to collect high resolution (100 Hz) acceleration data (g-force) allowing newer statistical/analytic processing possibilities for the assessment of PA/SB. As such, the science of calibration studies and corresponding analytical data processing has evolved to a current practice of semi-structured ADL performed in the laboratory, carried out for fixed-duration start-stop times across different activity domains (household, occupation, leisure, transportation). This permits the use of advanced machine learning metrics to reveal signal patterns to improve the prediction of free-living energy cost and activity type/domain.[42-62] Prior to our recent work **(R21HD080828)**, the scientific literature focusing on developing PA/SB accelerometer assessment methods for those with disability/ impairment employed older calibration approaches and analytics to algorithm development (i.e., linear regression or receiver operator characteristic binary outcome *"cut points',* and was starting to calibrate specific to disease condition (i.e. specific to Parkinson's disease[63, 64] or specific to multiple sclerosis[65, 66]). Our recent work (see preliminary studies) represents a scientific breakthrough, avoiding a disease-specific conundrum approach by classifying heterogeneous functionally limited populations into simple clusters for development of function cluster-specific accelerometer algorithms using advanced data processing.

2.a.4. New free-living accelerometer activity classification algorithms are necessary. To date most calibration studies developing accelerometer algorithms for the classification of activity energy cost and type have been developed on healthy people, and use laboratory protocols with fixed duration start-stop time ADLs.[43-45, 50, 60, 67-69] These calibrations in healthy populations, in concert with our recent methodological work specific to functional clustered populations, have permitted a proof-of-concept that machine-learning data processing approaches from wearable accelerometers result in improved estimates for predicting PA and SB. However, such artificial conditions (prescribed duration start-stop times) have facilitated high algorithm performance that has recently been shown to suffer significant degradation when applied to free-living conditions.[58, 70] Accelerometer signals from free-living conditions are substantially different from those obtained in the laboratory,[58] and there is a current understanding that free-living conditions are necessary to train accelerometer algorithms to improve the precision and accuracy of algorithm performance in real-world settings. The proposed studies will build upon our prior methodological work developing functionally clustered-specific accelerometer algorithms using advanced analytic data processing. In the current application, we will significantly extend calibration approaches using free-living activities to improve algorithm performance for activity energy cost and type prediction/classification in naturalistic settings.

2.b. INNOVATION.

In the current set of rigorous free-living best practice designed studies, we propose to refine and validate accelerometer algorithms using high-resolution 3-axis data from devices worn on the hip, wrist, and ankle specific to functionally clustered populations. These models based on standardized high-resolution wave form acceleration data will be applicable to other devices using 3-axis accelerometers, thus advancing other device approaches to improving free-living PA and SB assessment in diverse heterogeneous populations. Results will be easily applicable and will advance current national surveillance, epidemiological, and clinical research PA and SB objectives.

2.b.1. Clustering heterogeneous populations by functionality. Current practice applies accelerometer algorithms developed on healthy populations to those with functional movement limitations, providing erroneous estimates of PA and SB (as we demonstrate in our preliminary studies). The development of functionally clustered groups from the deployment of easily translatable clinical function tests will permit population groupings for use of accelerometer

algorithms calibrated specifically to level of functional impairment. Accelerometers represent a biomechanical movement sensor, predicated on physical movement. The current application innovatively aligns the design of a biomechanical movement sensor to the physical functional movement of populations to guide algorithm use.

2.b.2. Free-living behavioral activity energy cost prediction and activity type classification. High-resolution 3-axis accelerometer data have not been validated in free-living, non-scripted scenarios across diverse healthy and movement-limited populations. The current set of studies will use best practice calibration standards to validate energy cost, and to identify activity transition points (e.g., sit to stand or one activity to another), activity types, and domains along the continuum from SB to intense PA. The information will be applicable to broad populations varying in physical function classification and disability/impairment.

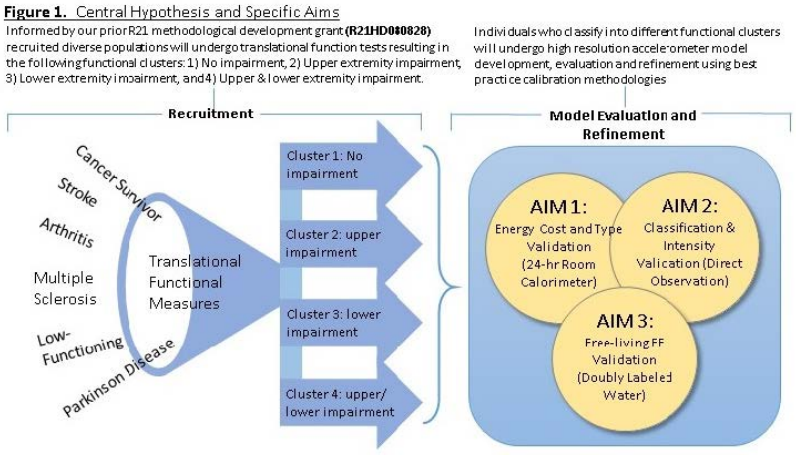
2.b.3. Optimizing precision and accuracy in free-living assessment. Our approach will have all individuals wear non-dominant hip, wrist, and ankle accelerometers to develop and refine individual or combined site placement algorithms for energy cost and activity type/domain prediction. Data will also provide an independent free-living dataset to validate current healthy population wrist algorithms for continued improvement of NHANES surveillance. Data will permit the researcher to optimize site placement accuracy and precision and further enable decisions on tradeoff between adding accelerometers to improve measurement versus feasibility/practicality of wearing more than one measurement device. Furthermore, results generated will also optimize accelerometer site placement for those with functional limitations.

2.b.4. Best-practice criterion for free-living studies. Recent research has exemplified the need to use free-living conditions to develop and refine accelerometer-advanced analytic algOrithMS.[43, 49, 52, 58, 59] The use of 24-hr indirect calorimetry chamber data will provide a highly controlled environment, but will permit criterion validity while participants carry out unstructured routines scientifically advancing current fixed duration start-stop duration laboratory calibrations. Our approach will provide a criterion to energy cost and activity type and also provide further understanding to advance classification of wear-time algorithms and sleep monitoring specific to functionally clustered groups across a 24-hr period. The methodological extension to use video-recorded direct observation in the field to validate physical function cluster-generated accelerometer algorithms is both innovative and essential to provide training data to improve natural environment PA/SB behavioral predictions. Fully synchronized free-living data will permit activity intensity, activity type, and activity transition to be simultaneously evaluated for improved algorithm performance. Differentiating free-living start and stop times and point of activity transition will advance activity type prediction and will significantly minimize accelerometer algorithm confusion.

3. APPROACH

3.a. Overview. This RO1 application represents three non-dependent, non-sequential aims designed to refine and validate machine-learned, body-worn, high-resolution accelerometer predictions of PA energy cost, activity types, and activity domains across a broad healthy and physically impaired heterogeneous population in naturalistic settings. A total of 380 participants (ages 18-100) with varying levels of physical function will be studied over a 5-year period. All individuals will undergo functional screening (upper and lower body) using translatable function tests to cluster into either a healthy movement cohort, lower body movement-impaired cohort, upper body movement-impaired cohort, or upper and lower body movement-impaired cohort. **Figure 1** below shows our central working hypothesis and aims. For Aim 1 carried out at the University of Colorado-Anschutz Medical Campus (UC-AMC), 90 individuals will participate in a

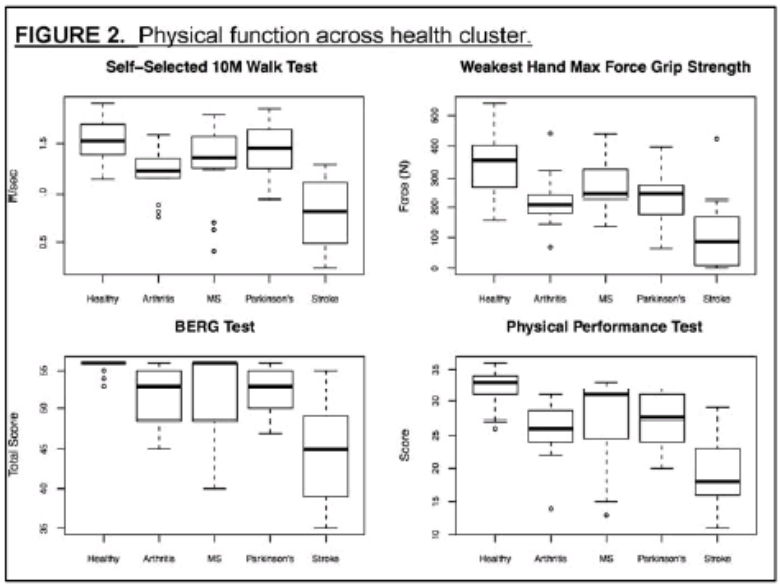
24-hr room calorimetry study, and will be asked to wear 3 accelerometers- one each on the non-dominant hip, wrist, and ankle. During the 24-hr calorimeter stay participants will be asked to complete a variety of exercises and daily living tasks, with no limit on duration or when the activity takes place (natural, but in a controlled setting). All 24-hr calorimeter data will provide criterion energy cost for functional clustered algorithm refinement and evaluation, and activity type and transition from video-recorded observation will be imported into a direct observation platform. For Aim 2 carried out at the University of Wisconsin-Milwaukee (UWM), 200 individuals will be monitored for a total of 3 hrs/day for 4 different days (total 12 hrs/person or a total of 2400 hrs, which is 144,000 min or 36,000 min of observations per hypothesized functional cluster). All sessions will be completed within a 14-day period. Days and times will be randomly selected and carried out in different environments (indoor, outdoor, work, home) and across different domains (occupational, transportation, leisure, household). All participants will be directly observed and video recorded during each 3-hr session with data imported into a direct observation platform, and all participants will also wear accelerometers placed on the non-dominant hip, wrist, and ankle. Direct video-recorded observation will provide a ground truth criterion for activity type/behavior, as well as a point estimate for PA intensity. For Aim 3 carried out at both UC-AMC and UWM, 90 participants will wear 3 accelerometers on the non-dominant hip, wrist, and ankle for a continuous 10-day observation period while measurements of PA-related energy expenditure are made via the doubly labeled water (DLW) technique.

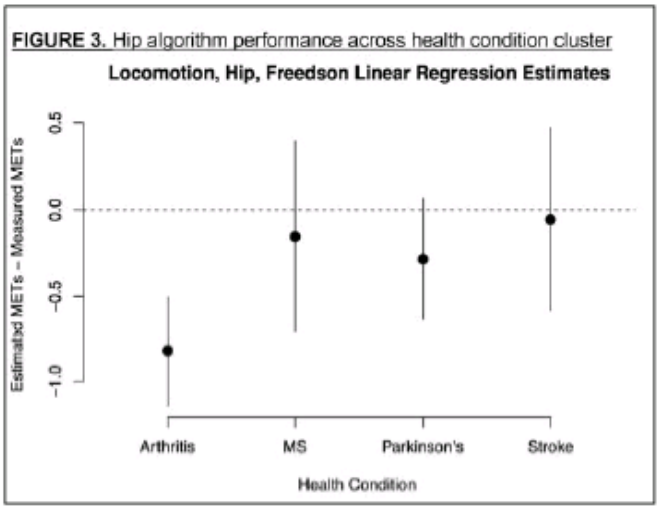


Reproducibility and Transparency: We propose to incorporate translatable functional tests to guide functional accelerometer algorithm choice, and developed accelerometer algorithms into freely available open-source software platforms. This will permit broad scientific deployment and advancement in an effort to minimize the scientific knowledge gap and practical deployment of advanced accelerometer algorithms to assess PA/SB of diverse heterogeneous populations.

3.b. Preliminary Studies. In July 2014, our research group (PI **Strath,** Co-Is **Staudenmayer, Hyngstrom, Swartz)** was funded **(R21HD080828)** to investigate the utility of accelerometer algorithm use in populations with disease and functional limitations. We set out to test a diverse heterogeneous population and measure physical function using an extensive range of available laboratory and field-based measures in an effort to cluster the populations into functional impairment groups, with accelerometer algorithm development to follow for each functionally clustered group using laboratory fixed start-stop time ADLs. We have tested n=122 individuals with a broad range of disease (cancer survivor, multiple sclerosis, Parkinson's disease, arthritis, stroke, and low physical functioning). This rigorous methodological work has led to the following conclusions:

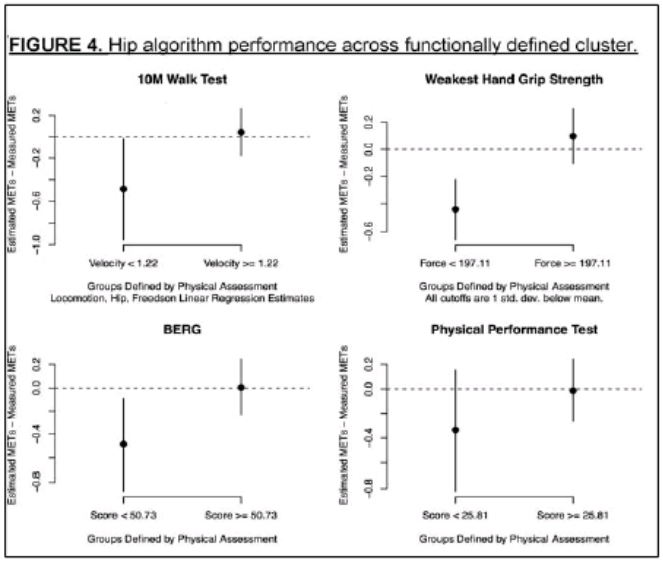
3.b.1. Physical function, as assessed by various methods, varies considerably within health condition. This is illustrated in **Figure 2** to the right, which shows the distributions of four physical function assessments in participants clustered by self-described health status. The overlap of the box-plot "whiskers" indicates that the body of the distributions (median +/- 1.5 IQR) overlap substantially between the healthy and diseased groups. Results were similar for other functional assessment methods, and these four assessments were chosen to illustrate two easily administered measurements (10 meter walk test and grip strength) and two assessments that require more time and expertise to administer/score (BERG balance test and Physical Performance Test (PPT)). Variability in physical function in individuals clustered by health status has an implication for the performance of methods to assess PA/SB with accelerometers. This can be clearly seen with the following preliminary results.





3.b.2. The performance of a widely used accelerometer method does not vary substantially across groups of individuals who are clustered by health condition. Figure 3 to the right illustrates this point by showing estimates of the bias (and 95% confidence intervals) of the Freedson *et al.* (1998)[71] method to estimate energy expenditure from a hip-worn accelerometer during locomotion activities. With the exception of the arthritis group the intervals span zero, indicating that estimates are not significantly biased for the disease groups of multiple sclerosis, Parkinson's, or stroke. Said in another way, disease cluster per se does not impact the performance of currently available accelerometer algorithms. This in turn led to our most important conclusion from our R21 simulated laboratory ADL methodological work:

3.b.3. The performance of previously published accelerometer-based methods to estimate aspects of physical activity/sedentary behavior in the laboratory *does* vary substantially across groups of individuals who are clustered by physical function. Figure 4 (to the right) **and Figure 5** below illustrate this essential point. **Figure 4** assesses the performance of the Freedson *et al.*



(1998)[71] hip accelerometer method to estimate energy expenditure during locomotion activities on participants who are clustered according to the four assessments of physical function listed above in Figure 2. In all cases, the functional-clustered groups are formed by estimating a cutoff at one standard deviation (SD) below the mean value for that functional test in the study population. The actual functional cutoffs are listed in the figure. Sample mean was used in conjunction with population normative data when available. The established method using the Freedson *et aL* (1998)[71] hip algorithm significantly underestimates energy expenditure in the functionally limited clusters (individuals 1SD below the mean), thereby indicating erroneous estimates when applying "healthy" population accelerometer algorithms to those with functional limitations.

Figure 5 (to the right) repeats the approach of **Figure 4**, but assesses the performance of recently developed machine learning methods (Staudenmayer *et al.,* 2015)[59] to classify activity types (% correct) from a wrist-worn accelerometer. The method to classify locomotion (versus not) and the method to classify sedentary time (versus not) both perform worse (up to a 10% decrement in activity type classification) in the functionally limited population (1 SD below the mean for each respective functional measure). This further highlights that whether traditional accelerometer hip or newer accelerometer wrist algorithms are used, accuracy to estimate energy cost or activity type is compromised when applied to those with functional limitations. This necessitates filling a scientific knowledge gap by developing accelerometer algorithms for those with functional limitations using best practice, field-based calibration techniques.

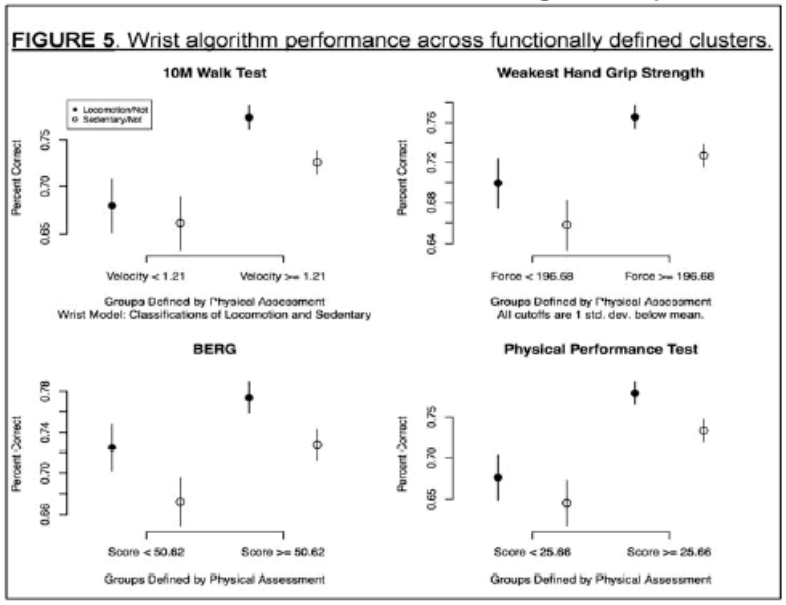


FIGURE 5. Wrist algorithm performance across functionally defined clusters.

Table 1. Probability clustering of physical function measures across diverse populations into above and one standard deviation below mean.

	DGI_Total	BERG_Total	PPT_Total	TEN_M_Velocity	BOX_AVG	Lower.limb.max.force	Hand.max.force	Lower.limb.force.stead	Hand.force.steadiness	MMT_Upper	MMT_Lower	PPT_Upper	PPT_Lower
DGI_Total	1.00	0.93	0.88	**0.93**	0.84	0.81	**0.84**	0.71	0.85	0.88	0.90	0.81	0.94
BERG_Total		1.00	0.90	**0.95**	0.85	0.81	**0.87**	0.74	0.87	0.89	0.91	0.84	0.96
PPT_Total			1.00	**0.93**	0.87	0.82	**0.89**	0.71	0.89	0.89	0.89	0.84	0.94
TEN_M_Velocity				**1.00**	**0.83**	**0.83**	**0.87**	**0.72**	**0.87**	**0.87**	**0.89**	**0.83**	**0.99**
BOX_AVG					1.00	0.82	**0.83**	0.74	0.86	0.88	0.86	0.85	0.84
Lower.limb.max.force						1.00	**0.87**	0.69	0.81	0.81	0.83	0.83	0.82
Hand.max.force							**1.00**	**0.74**	**0.92**	**0.92**	**0.88**	**0.86**	**0.86**
Lower.limb.force.steadiness								1.00	0.82	0.78	0.78	0.78	0.71
Hand.force.steadiness									1.00	0.94	0.90	0.88	0.86
MMT_Upper										1.00	0.92	0.86	0.86
MMT_Lower											1.00	0.84	0.90
PPT_Upper												1.00	0.83
PPT_Lower													1.00

DGI=dynamic gait index, BERG=Berg balance test; PPT=physical performance test; TEN-M-Velocity=10meter walk test average velocity; BOX=Box and Block test; All force tests were maximal voluntary contraction force; force steadiness = coefficient of variation around holding a static contraction at 25% maximal; MMT=manual muscle test.

The four functional assessments included in these examples identified similar clusters of individuals. This is illustrated for these assessments and others in **Table 1** to the right. **Table 1** shows the percent of participants put in the same high or low function cluster by each assessment method. This result suggests that for the specific purpose of determining accelerometer algorithm applicability, simple and field translatable tests such as the 10 meter walk test and a maximal

grip strength test could be used to discriminate between a) normal functioning, b) lower-impairment, c) upper-limb impairment, or d) lower- and upper-limb impairment.

Summary of Preliminary Studies: These studies demonstrate that accelerometer algorithms should be developed and validated on physical movement patterns, *not* disease status. Our team has evaluated translatable functional measures to successfully cluster physical function and has further highlighted the current erroneous estimates of available hip or wrist accelerometer algorithms for estimating energy cost or activity type across functionally clustered groups.

There is a critical knowledge gap concerning how to assess PA/SB in populations with disability and movement limitations, which lays the essential foundation for the current RO1 application: to develop, evaluate, and refine high-resolution accelerometer machine learning algorithms from innovative free-living activities specific to functionally clustered populations.

3.c. Investigative team. Our research team includes a unique combination of investigators with complementary expertise in exercise science (Strath, Swartz, Melanson), statistics and mathematics (Staudenmayer), and physical medicine and rehabilitation (Hyngstrom). Collectively this team has a combined 40-year history of working in the area of PA assessment (Strath, Swartz, Staudenmayer, Melanson) and are leaders in the field for analytical advancement in modeling high-resolution accelerometer data (Staudenmayer). Investigators for the proposed studies work well as a team, are highly productive, and just recently carried out prior methodological work **(R21HD080828)** in accordance with timelines indicated. Our team is uniquely qualified to carry out the aims as outlined in this application and will ensure success of the proposed work.

3.d. AIM 1. Evaluate and refine machine-learned algorithms to predict energy cost and discriminate activity type during a 24-hr respiratory calorimeter stay. Hypothesis: Analytical models that are derived on physical function clusters using machine learning methods will produce more accurate and precise estimates of PA intensity and type classification compared with models developed using healthy populations.

3.d.1. Participants: Screening and recruitment. Ninety individuals will be tested ranging in age from 18-100 years, with varying levels of functional ability. Individuals who meet inclusion and exclusion criteria **(Table 2** at right) will undergo a physical functionality screen. Screening tests as outlined in **Table 3** below will be used to place participants into respective normal, lower, upper, or combined lower and upper functional impairment groups. These translatable screening tests to guide functional group placement and ultimately a functionally matched accelerometer algorithm are based upon our prior work **(R21HD080828)** and in conjunction with preliminary studies presented herein. We have elected to still conduct the physical performance test as a measure to solidify our translatable functional measures to cluster successfully. The current study sample will be stratified by age and gender, and we will continually evaluate recruitment across

Table 2. Inclusion and Exclusion Criteria for Aims 1-3

Inclusion Criteria
• Age greater than or equal to 21 and less than or equal to 100 years of age
• Able to ambulate
Exclusionary Criteria
• Wheelchair reliant
• Assistive walking device reliant
• Uncontrolled hypertension with a blood pressure greater than 160/100 mmHg at the initial visit
• Known history of cognitive impairment or inability to follow study procedures
• Metabolic altering medications

functional groups to maintain equal cell/cluster distribution. Sample size estimates are presented below. We will draw from a multitude of ambulatory clinical patient groups including but not limited to the following: Cancer survivor, stroke, multiple sclerosis, Parkinson's disease, arthritis, chronic pain, low-functioning, general population. Populations will be drawn and screened in local clinics and from institutes specific to functional/disease state, as well as through general advertisements and mailings. Our investigative team has a strong working relationship with all population subgroups, clinic and disease institutes, and we have successfully screened and recruited hundreds from similar populations in prior work **(R21HD080828).** In conjunction with accompanying letters of support and past experience we do not anticipate any difficulty with participant recruitment or retention.

Table 3. Functional Screening Measures of Upper/Lower Impairment.

Test Measure	Abilities/Constructs Assessed
1. Upper Extremity (repeated bilaterally)	
Hand Maximal Dynamometry[94]	Hand maximal strength/strength
2. Lower Extremity	
Self-Selected and Maximal Walking Speed[95]	Gait speed determined over 10 meters on a level surface/function
3. Combined Upper/Lower Extremity	
Physical Performance Test[96]	9 item inventory tasking everyday activities/function, strength, and coordination

3.d.2. Data collection and measures. Following screening and enrollment, participants will make two visits to the UC-AMC. Visit 1 will consist of written and witnessed informed consent, a medical history to confirm eligibility, and anthropometric measures including height, weight, body girth, and resting blood pressure.

Individuals will then undergo measures of resting metabolic rate (RMR), and 3-compartment body composition analysis using dual-energy x-ray absorptiometry (DXA). At the end of Visit 1, participants will be orientated to the room calorimeter by study staff. Measures are presented and described in Table 4 (below). Participants will then return to the UC-AMC for Visit 2, which will consist of a 24-hr room calorimeter stay while wearing accelerometers affixed to their non-dominant hip, wrist, and ankle via an elastic belt and Velcro straps.

3.d.2.1. Room Respiration Calorimetry. The room calorimeter is a 12'x12' ventilated chamber where a subject may reside comfortably for long periods of time unencumbered by typical portable gas collection devices. Total energy expenditure (TEE) oxygen (02) and carbon dioxide (CO2) concentrations will be measured continuously using a fuel cell-based dual channel 02 analyzer (FC-2 Oxzilla, Sable Systems International, Las Vegas, NV) and two infrared CO2 analyzers (CA-10 CO2 analyzers, Sable Systems International, Las Vegas, NV), as previously described.[72] 02 consumption (V02) and CO2 production (VCO2) will be calculated in 1-min intervals using flow rate and the differences in CO2 and 02 concentrations between entering and exiting air, and minute energy expenditure (EE) will be calculated using the equations of Jequier *et al.*[73] TEE will be obtained by summing minute values for EE. PA energy expenditure will be calculated as the difference between 0.9 TEE and RMR, assuming the thermic effect of food is equal to 0.1 TEE.

The calorimetry protocol will include an awake and a sleep period. While inside the chamber all participants will be asked to complete a list of activities from within different domains of activity. Activities will reflect different types, domains, and intensities (from sedentary low-energy level to more intense high-energy level). Activities will not be scripted or scheduled, but rather represent a check list of items that must be completed for a minimum period of time (5 mins or more) during the awake portion of the calorimeter stay. This will permit an evaluation of activity and transition to be monitored. An example list of activities is provided below:

- **Sedentary behaviors:** Reading, writing, watching television, card/board games.
- **Household tasks:** Sweeping the floor, dusting, vacuuming, cleaning windows.
- **Occupational tasks:** Computer work, moving items.
- **Transportation:** Treadmill walking at different speeds, less than normal pace, normal pace, and faster than normal pace.
- **Leisure:** Walking/running/cycling, light calisthenics, light resistance dumbbells, yoga, tai chi.

Participants will report to the chamber at 0800 on the day of observation and stay for 24 hrs. Participants will be fed three standard meals served at 0900, 1300, and 1800, and a morning and evening snack. Meal composition will not be strictly controlled as this is not an energy balance study - participants may choose from

Table 4. Physiological, accelerometer measures: Specific Aim 1-3.

Test Measure	Description
Anthropometric	Height and weight will be assessed using a stadiometer and physician scale. Body mass index (BMI) will be calculated as a ratio of body weight to height squared. Girth measurements will be assessed at the waist and the hips (maximal circumference of the buttocks above the gluteal fold) with tension-fitted measurintape.
RestingMetabolicRate (RMR)	RMR will be assessed utilizing the flow-through hood technique (ParvoMedics TrueOne 2400). We have previously validated this system forthe measurement of oxygen uptake.[97] Each individual will be asked to strictly com ply to requirements for an RMR test. Individuals will be assessed after laying supine for a minimum of 30 minutes in a thermoneutral environment. Min-by-min data will be analyzed to determine RMR during a steady state period of 10 min or more.
Dual energy x-ray absorptiometry (DXA)	DXA.assesses body com posit on in 3 compartments: mineral-free lean, fat, and bone mineral content. Data will be provided on total body lean tissue, total body fat mass and percent body fat, bone mineral content and density.[93]
Accelerometer	The Acti graph Gt3X+BT (Actigraph LLC, Pensacola, FL) will be attached to the non-dominant hip, wrist and ankle to assess movement during each study Aim. The device uses a micro-electro-mechanical system (HEMS) based accelerometer with dynamic range of +/- 6 G. The acceleration data are sampled by a 12-bit analog digital convertor and will be set at a range of 100Hz and will be stored in high-resolution format.

a menu of food choice. Water is permitted at any time during the stay. Participants will leave the chamber the following day at 0700. Temperature inside the calorimeter will be maintained at 72°F (22°C). During the course of the stay all activities will be video-taped. Video recordings will

be stored and sent for analysis at the UWM site (**see Data Sharing and Management Plan**) at which time recordings will be directly imported into a direct observation analysis system (Noldus Information Technology, Netherlands — further discussed under Aim 2). All data collection tools (minute-by-minute room calorimeter, accelerometers, and video recordings) will be time synchronized.

3.d.3. Aim 1: Analysis plan: **This analysis plan applies to all three Aims in the current application.** There are two types of statistical analyses in this project: Machine learning model development and model comparison to test the hypotheses in the Aims. We discuss the machine learning model development first, and we specify how we will evaluate the hypotheses at the end of this sub-section.

Machine learning model development: These studies will generate several streams of best practice calibration data from which we can develop statistical models to relate accelerometer signals (covariates) and aspects of PA/SB (responses) and customize those models specific to discriminate clusters of functional impairment (none, lower-limb impairment, upper-limb impairment, and lower- and upper-limb impairment). The data in the current application, which will provide innovative and essential free-living naturalistic setting observations, will be further augmented with laboratory calibration data (from block randomized 10 simulated ADLs) from the previous R21 project (n=122).

We propose to implement those models as in previously successful projects,[59, 74] and we refer to that work for a more detailed and technical discussion of analytical approach. An overview of the approach is as follows. We define the covariates and responses; we describe the types of models we will use; and we overview how we propose to evaluate model performance.

Table 5. Features used to summarize the accelerometer signals in an interval.

Feature	Definition
Mean of vector magnitude (vm) (mvm)	This statistic is the sample mean of vm in interval
SD of vector magnitude (sdvm)	This statistic is the SD of vm in interval.
Percentage of the power of the vector magnitude that is in 0.6-2.6 Hz (p625)	Estimate a spectral density of vm using a fast Fourier transform. Compute the modulus corresponding to each frequency. This statistic is the sum of those moduli divided by the sum of the moduli at each frequency.
Dominant frequency of vector magnitude (df)	This statistic is the frequency that corresponds to the largest modulus.
Fraction of power in dominant frequency (fpdf)	This statistic is the modulus of the dominant frequency divided by the sum of the moduli at each frequency.
Mean angle of acceleration relative to vertical on the device (mangle)	Compute the angle = 90arcsin(x/vm)/(pi/2). This statistic is the sample mean of those angles in the interval.
SD of the angle of acceleration relative to vertical on the device (sdangle)	This statistic is the sample SD of the angles in the interval.

We will use flexible regression models to fit a model to the calibration dataset, and models relate features of the accelerometer signal and measures of physical limitation/impairment (covariates) to aspects of concurrent PA/SB (response). **Table 5** above gives an example of features that we will use. These features are typically computed from 10- or 15-second intervals

of accelerometer signal, and most of these features have been used in previous work, [45, 57, 59, 68, 74] which also describes why they have been chosen.

The aspects of PA/SB that we will predict fall into two general types: continuous response and categorical response. Metabolic equivalents (METs — energy cost divided by measured RMR) will represent the continuous response. The direct observation and the Compendium values, the 24-hr calorimetry data, and results from the DLW will be used to develop the calibration dataset. We will also develop models with categorical responses including level of activity (sedentary, light, moderate, and vigorous) and type of activity. We will consider a number of different approaches to describe type of activity in an interval such as sedentary or not, and more detailed descriptions of activities, locomotion or not, sports or not, and multi-category descriptions: sedentary, locomotion, vigorous sports, intermittent movement. In these cases also, direct observation will give the value of these responses for the calibration dataset. All models will be evaluated for sex differences.

Next, we describe the models we will use. The general class of models that we will use is known as statistical (or machine) learning models. While these can be viewed as regression models because they use inputs (the features or covariates) to predict the response, they extend linear (or generalized linear) regression models by not assuming rigid parametric relations between the covariates and the mean response. Instead, the machine learning approaches use data to flexibly learn that regression relationship under minimal assumptions. One danger in using statistical learning approaches is that they can overfit both the signal and stochastic errors. In that case, the model would fit one particular dataset very well, but they would not perform well when applied to a new dataset. To avoid this, the estimation of statistical learning models includes a regularization step to reduce the flexibility of the model. More detail about statistical learning can be found in other sources.[75, 76] We will use at least three specific statistical learning models: neural networks, support vector machines, and random forests. These approaches have been used to estimate PA from accelerometers previously.[45, 57, 59, 68, 74]

Finally, we will evaluate and compare model performance using bias, standard error, and root mean squared error for continuous estimate and percent correct/confusion matrices for categorical estimates. The categorical responses will also be evaluated with activity-specific sensitivity and specificity. All results will be cross-validated so that the same subject's data are never used to both build and test a model. When appropriate, we will use multiple comparison correction procedures and mixed models to address repeated measures. Mixed models will be used to test the hypotheses and compare model performance across functionally clustered groups (i.e. healthy population algorithm on functionally limited group(s)). Statistical analyses, including the preprocessing to compute statistical summary inputs and specific implementation of the statistical learning methods, will be performed using the free R-software.

3.d.3.1. Sample Size Estimation: Standard sample size estimations are difficult when determining training data sets. In this analysis specific to Aim 1 the goal is to have at least as many participants and observations as have been used in other previously successful calibration studies in both adults[43, 48, 87, 59, 68, 74] and children.[62, 77, 78] To establish a sample size to give adequate power to reliably detect practical significant differences in our hypotheses, we will consider the estimation and testing of both METs as a continuous variable and classification of activity as a percentage. To guide sample size estimations we consider the work of Lyden *et a1.*[79], senior author Melanson (Co-I on this application). With a 10% desired effect size and a subject-to-subject standard deviation of less than 15% of mean energy expenditure estimations, we estimate a sample size of n=20 would be sufficient to yield 80% power. This is less than our total sample size (n=90), indicating that we will be able to reliably test our hypothesis for Aim 1

on functionally clustered groups as well as testing our hypothesis on the whole sample. To evaluate classification accuracy, we will test functionally clustered specific models compared with 80% classification accuracy. Using a sample size of n=90, alpha=0.05, and a true classification accuracy of 50% (conservative estimate), we would have 80% power to detect 10.5% deviation from a threshold of 80%. Given this, 90 subjects will be recruited for Aim 1.

3.e. AIM 2. Validate machine-learned accelerometer algorithms with field-derived, video-recorded direct observation. Hypothesis: Physical function-clustered, specific machine-learned accelerometer algorithms will demonstrate more accurate and precise PA intensity and activity type classification during free-living compared with healthy population derived accelerometer algorithms applied to diverse populations.

3.e.1. Participants: Screening and recruitment. Two hundred individuals will be tested ranging in age from 18-100 years, with varying levels of functional ability. Individuals who meet inclusion and exclusion criteria (see **Table 2)** will undergo a physical functionality screen. Similar to Aim 1, screening tests as outlined in **Table 3** will be used to place participants into respective normal, lower-, upper- or combined lower- and upper-impairment functional groups with screening taking place at surrounding clinics. The study sample will be stratified by age and gender, and we will continually evaluate recruitment across functional groups to maintain equal cell/cluster distribution. Sample size estimates are presented below. We will draw from a multitude of ambulatory clinical patient groups including but not limited to the following: Cancer survivor, stroke, multiple sclerosis, Parkinson's disease, arthritis, chronic pain, low-functioning, and general population. Populations will be drawn and screened in local clinics and from institutes specific to disease state, as well as through general advertisements and mailings. Similar to Aim 1, we do not anticipate any difficulty with participant recruitment or retention.

3.e.2. Data collection and measures. Following screening and enrollment participants will be contacted and our team will meet them at their place of residence. This initial visit will consist of written and witnessed informed consent, a medical history, and anthropometric measures including height, weight, and resting heart rate and blood pressure. Measures are described in **Table 4.** At this time participants will be fitted with 3 Actigraph Gt3X+BT accelerometers affixed to their non-dominant hip, wrist, and ankle with an elastic belt and wrist/ankle Velcro straps. Each person will be asked to carry out their typical daily routine activities for 3 hrs. During this time research staff will observe from afar, and unobtrusively video record free-living activity behavior. This will be repeated for a total of 4 direct observations, each lasting 3 hrs (total 12 hrs direct observation) and each observation period being at a different time interval per day to increase the representativeness of daily observations (i.e. observations span morning, afternoon, evening). Individuals will be asked to submit to the research staff a potential list of things to be accomplished during the 3-hr monitoring period — for instance, errands, housework, office work, yard work. We have employed this approach prior with great success[80] so as to avoid an observation period only consisting of one activity type, i.e. reading a newspaper for 3 hrs.

3.e.2.1. Direct Observation. Our study team has conducted extension pilot testing of recordable devices from which to collect direct observation data. We have elected to use the GoPro HERO camera that offers a wide angle recording view and LCD screen to ensure research staff capture images successfully. Digital format from the GoPro HERO can be directly uploaded into our Direction Observation Platform.

3.e.2.2. Direct Observation Coding Platform. We will utilize the Noldus™ direct observation customizable software (Noldus Information Technology, Netherlands) to code all uploaded

video. Our research group has successfully utilized this software to assess the in-laboratory and free-living PA profiles of older adults.[80] Other research groups have also shown the efficacy of this approach to assess PA/SB behaviors in different settings.[81] Utilizing the Noldus™ platform we will upload all video footage, time synchronize all accelerometer data with video footage (using the NoldusTM Data Module), and have trained coders pass through the footage to record activity type/domain, activity duration, and PA intensity. Our approach will use focal sampling, and record of each new activity undertaken. The metabolic cost (in 0.5 metabolic cost increments) can be assessed for directly observed behavior by examining each activity domain, type, intensity, and point estimate. We obtain the metabolic cost using drop-down menus that utilize domain selection (i.e. household), followed by activity type (i.e. cooking - with activity types guided by the Compendium of Physical Activities[82]), followed by intensity category. Activity duration is deduced from the time lapse in change of activity sequence and represents a running second-by-second clock. We will also have trained coders perform a second pass through video footage to isolate activity transition time (i.e. transition from cooking to eating or sitting to walking). <u>These data can only be obtained from video-recorded isolation and will be used for algorithm training purposes on transition time to limit actual activity confusion.</u>

Our direct observation coding staff will undergo extensive training on laboratory master video-recorded files that have already been pre-coded. Our research group[80] and others[81] have successfully trained coding staff in this manner for many observation protocols, with high intra-rater reliability across activity type and intensity selection. We will employ the same methodologically rigorous approaches necessitating intraclass correlation coefficient values above 0.90 across coders. Further, all uploaded video files will be coded in duplicate to further minimize error and bias. The volume of data to be coded will be continuously on-going, so we will require all staff to re-certify training and inter-rater reliability every 6 months of active data collection/coding.

<u>3.e.3. Aim 2: Analysis plan.</u> The analysis plan to guide accelerometer model development, refinement, and testing will be **<u>the same as described in Aim 1.</u>**

3.e.3.1. Sample Size Estimation: In addition to ensuring that hypothesis tests are adequately powered, the sample size also needs to be large and diverse enough to train the machine learning models. Standard power and sample size analyses are not available for this purpose; instead we will ensure that we have at least as many subjects (and observations per subject) as have been used by previous successful studies. Specifically, we consider three recent studies: a free-living study in adults by Lyden *et al.*[81] and two lab-based studies of adults (Strath *et al.*[43]) and youth (Trost *et al*.[82]). The free-living study in adults found that n=32 subjects (6 hrs each) were adequate to reliably train machine learning models. The lab-based study in adults used n=99 (33 per age group demarcation) and the lab-based study in youth used n=100, aged 5-15 yrs (9 per age group demarcation). Collectively, this gives us confidence that 200 subjects, measured for 12 hrs each, across 4 functionally clustered groups, will yield reliably trained machine learning models.

3.f. AIM 3. Validate machine-learned algorithms using the doubly labeled water technique. <u>Hypothesis:</u> Physical function-clustered specific machine-learned accelerometer algorithms will produce more accurate and precise estimates of PA energy expenditure (PAEE) during free living compared with healthy population derived accelerometer algorithms applied to diverse populations.

<u>3.f.1. Participants: Screening and recruitment.</u> Ninety individuals will be tested ranging in age from 18-100 years, with varying levels of functional ability. Individuals who meet inclusion

and exclusion criteria (see **Table 2)** will undergo a physical functionality screen. Similar to Aim 1 and 2, screening tests as outline in **Table 3** will be used to place participants into respective normal-, upper-, lower-, or combined upper- and lower-body impairment functional groups with screening taking place at surrounding clinics. The study sample will be stratified by age and gender, and we will continually evaluate recruitment across functional groups to maintain equal cell/cluster distribution. Sample size estimates are presented below. We will draw from a multitude of ambulatory clinical patient groups including but not limited to the following: Cancer survivor, stroke, multiple sclerosis, Parkinson's disease, arthritis, chronic pain, low-functioning, and general population. Populations will be drawn and screened in local clinics and from institutes specific to disease state, as well as through general advertisements and mailings. Similar to Aim 1 and 2, we do not anticipate any difficulty with participant recruitment/retention.

3.f.2. Data collection and measures. This Study Aim will be carried out at both the UC-AMC (n=45) and the UWM (n=45) site. Visit 1 will consist of written and witnessed informed consent, a medical history to confirm eligibility, and anthropometric measures including height, weight, resting heart rate, and resting blood pressure. Individuals will then undergo measures of resting metabolic rate, and 3-compartment body composition analysis using dual-energy x-ray absorptiometry (DXA) **(See Table 4).** At the end of Visit 1, individuals will undergo procedures for the DLW measurement, as below, and will be fitted with 3 Actigraph Gt3X+BT accelerometers affixed to their non-dominant hip, wrist, and ankle with an elastic belt and wrist/ankle Velcro straps. Each person will be asked to carry out their typical daily routine activities for 10 days then return to the respective laboratories for conclusion of the DLW procedures and to return accelerometers.

3.f.2.1. Physical activity energy expenditure & total energy expenditure: TEE will be measured over 10 days using DLW. Prior to dosing, a baseline urine sample will be obtained for determination of background enrichments of $^{2}H_2$ and ^{18}O. Participants will then consume an oral dose of water containing 1.8 g/kg total body water (TBW, estimated as 73% of FFM) of 10 atom percent excess (APE) ^{18}O and 0.12 g/kg TBW of 99.9 APE ^{2}H.[83] Urine samples will be obtained 4 and 5 hrs, after dosing, and plasma from a 5 mL blood sample will be obtained at 5 hrs. On day 10, participants will be instructed to discard their first urine void of the day. They will then report to the laboratory, and the second and third urine voids of the day will be collected, along with a plasma sample. Sample aliquots (4 mL) will be frozen at -10˚C until analysis. UWM samples will be shipped using standard methods to UC-AMC for analysis.

Frozen urine samples will be thawed and prepared by centrifugation and analyzed for ^{18}O and $^{2}H_2$ enrichment by Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS, Los Gatos Research Inc, Mountain View CA), as previously described.[84] Briefly, OA-ICOS uses a laser-based methodology and poses advantages of increased throughput and thus faster analysis of samples as compared to traditional isotope ratio mass spectrometry (IRMS). We have recently reported[85] that TEE (i.e. CO2 production) measured using OA-ICOS over 7 days is valid compared to simultaneous whole-room indirect calorimetry and IRMS. Data are analyzed using commercially available Post Analysis Software (Los Gatos Research Inc, Version 2.2.0.12), which utilizes inter-run standard measurements to automatically calibrate isotope measurements. Samples are run in duplicate, and repeated if the SD exceeds 3%. Dilution spaces for ^{2}H and ^{18}O will be calculated from the baseline samples according to the methods of Coward.[86] Total body water will be calculated as the average of dilution spaces of ^{2}H and ^{18}O after correction for isotopic exchange with other body pools.[87] CO2 production rate will be calculated using a modification[88] of the original two-point equation of Schoeller et al.[83] TEE will be calculated using the equation of Weir,[89] assuming a respiratory quotient of 0.86, and averaged over 10d. PAEE is the difference between TEE and REE, adjusted for the estimated

thermic effect of feeding (TEF).[90] In the UC-AMC lab, the average within-subject CV for repeat measures of TEE is 6.1±3.8%, slightly better than that reported (7.8%) across multiple labs.[91]

3.f.3. Aim 3: Analysis plan. The analysis plan to guide accelerometer model development, refinement, and testing will be **the same as described in Aim 1.**

3.f.3.1. Sample Size Estimation: The power and sample size were estimated based on the primary outcome of interest being the difference in PAEE obtained from the functionally clustered specific accelerometer algorithms and the DLW technique. A variety of models exist in the literature predicting PAEE from different activity monitoring types, worn in different locations, and using different unit outcomes. Two recent studies report on both healthy (Colbert *et al.* 2011[92], n=56 mean age 74.7 ± 6.5 yrs) and diseased populations (Rabinovich *et al.* 2013[93], n=39, mean age 69 ± 6.6 yrs) using the Actigraph accelerometer both worn on the waist, with Colbert *et a1.*[92] reporting relative to uniaxial count output and Rabinovich *et a1.*[93] reporting relative to vector magnitude output. Study models incorporated crude accelerometer output predictions of PAEE and also reported additional models with added predictor variables (age, gender, body weight). The shared variance for accelerometer model estimates and DLW range from R^2=0.40 to R^2=0.46. Thus with a power of 0.90 and an alpha of 0.05, a minimum shared variance of 0.43 can be detected with approximately 40 subjects. Given the breadth of subjects across functional clusters we have elected to initially double this sample estimate. Given this conservative estimate, and assuming a potential dropout/data loss rate of approximately 12%, 90 subjects will be recruited for Aim 3.

3.g. Potential problems and solutions. It is possible that our clustering for functional impairment will not adequately distinguish overall movement limitations. Given the strength of our pilot work and prior funded work **(R21HD080828)** this is highly unlikely. We are prepared to reorganize classifications if need be through alternative clustering approaches, for example examining gradations of distance from the mean to distinguish the poor functioning from the very poor. We have also proposed to collect additional field measures of physical function. Pilot work would indicate this is not needed, but if reclassifying function is necessary this will give additional functional data to reorganize in a meaningful way.

All major resources for the proposed studies are readily available at UC-AMC, UWM, and UMass Amherst. Drs. Strath, Staudenmayer, Melanson, and Swartz have extensive experience using wearable monitors, room calorimeters, direct observation, and DLW in adults. No problems are foreseen with data collection.

There will be a relatively large volume of data produced from the study Aims proposed. Data management and organization will be essential for success. Our team has successfully managed this previously **(R21HD080828)** across 2 sites. In the proposed application we add a third site (UC-AMC). We will take great care with study organization, flow, and data management as mentioned. We have further elicited the assistance of **SPADES,** multi-tier cloud infrastructure to store, visualize, and analyze data. Further details are provided within the **Data Sharing and Management Plan,** and accompanying letter of support from Dr. Fadh Albinali.

3.h. Timelines and Deliverables. Earliest project start date is 4/1/17. The full study timeline is provided in **Table 6** below. The first 8 months will be devoted to developing and finalizing rigorous data collection protocols and data management/sharing plans. The proposed studies consist of data collection across 2 sites, with simultaneous collection of Aim 1 and 2 data, and Aim 3 commencing in late year 3. Protocols are designed to continually provide training data to constantly evaluate and refine functionally clustered specific accelerometer algorithms, so Aims

are not dependent upon one another and can run simultaneously. Due to chamber time use we will be able to test 4/5 participants per month in Aim 1. For Aim 2, which will involve 4 different days of observation for 3 hrs/day, we anticipate being able to recruit and test approximately 8 subjects/month. Aim 3 will constitute each site recruiting/testing 45 subjects each, at a rate of 4/5 per month. Aims 1 and 2 will specifically take considerable personnel time to code videos, performed in duplicate for quality control.

Table 6: Proposed study timeline.

Study Year	Year 1			Year 2			Year 3			Year 4			Year 5		
Study Month	**1-4**	**5-8**	**9-12**	**1-4**	**5-8**	**9-12**	**1-4**	**5-8**	**9-12**	**1-4**	**5-8**	**9-12**	**1-4**	**5-8**	**9-12**
Site protocol & data plan	X	X													
Site audio visual meeting	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Site in person meeting	X		X	X			X			X			X		
UC-AMC Site															
Site Training	X	X		X			X			X					
Aim 1 Recruit/test			X	X	X	X	X	X	X	X					
Aim 3 Recruit/test										X	X	X	X		
Data Analysis						X	X	X	X	X	X	X	X	X	
UWM Site															
Site Training	X	X		X		X		X		X		X			
Aim 2 Recruit/test			X	X	X	X	X	X	X	X	X				
Aim 3 Recruit/test										X	X	X	X		
Data Analysis						X	X	X	X	X	X	X	X	X	
Dissemination							X	X	X	X	X	X	X	X	X
UMass Site															
Data Share/analysis			X	X	X	X	X	X	X	X	X	X	X	X	X

3.i. Proposed impact and future directions. Results will provide evidence that diverse populations (healthy and those with disabling/impairment conditions) can be meaningfully categorized with translatable functional screening tests and function clustered-specific accelerometer algorithms to accurately and precisely assess free-living PA and SB. Given the prevalence of the population (1 in 5 adults) who report a limitation to carrying out ADLs and IADLs, this is an essential area of scientific enquiry.

The breadth of best-practice calibration approaches used and the heterogenous populations studied and the data generated will provide an unprecedented opportunity to explore and extend open-source analytical and data mining methods. Data generated will also provide ample opportunities to ask ancillary scientific questions proposed by the investigators or other interested parties.

The results from the proposed studies using best practice free-living calibration protocols will provide for the first time the basis to accurately and precisely assess the PA behavior of diverse populations allowing for clear delineations of PA/SB prevalence rates, effectiveness of behavioral-based and therapeutic interventions, and dose-response PA/SB relationships to prevent and manage disease and impairment conditions.

4. PROTECTION OF HUMAN SUBJECTS

4.a. Risks to Human Subjects

4.a.1. Human Subjects' Involvement and Characteristics. Participation of subjects for Study Aims 1, 2 and 3 will involve the measurement of functionality using typical clinical measures of strength, mobility and gait speed, measures of height, weight, resting heart rate and blood pressure. A detailed health history will be conducted to guide stringent inclusion and exclusion criteria as outlined in the Research Strategy for all study Aims. Individuals will be drawn from local surrounding clinics and the general community, and will likely include individuals with stroke, multiple sclerosis, Parkinson's Disease, arthritis as well as those without documented disease/disability. All recruitment will include both men and women. Selection criteria include individuals between the ages of 21 and 100 years. A total of 90 subjects will participant in Study Aim 1. These participants will also undergo measures of resting metabolic rate (RMR) via flow-through indirect calorimetry, 3-compartment body composition using dual energy x-ray absorptiometry (DXA), and be asked to wear 3 affixed accelerometers to the hip, wrist and ankle while spending 24-hrs in a whole room calorimeter while performing selected exercises and activities of daily living. This is a 12'x12' room that has a large window, from which minute by minute oxygen consumption can be determined. For Study Aim 2, 200 participants will be asked to wear 3 accelerometers affixed to the hip, wrist and ankle for 3 hours each of 4 days. During this time all activities will be videotaped for purposes of direct observation determination of physical activity type, domain and energy cost. For the last study, Study Aim 3, 90 subjects will undergo measures of RMR, DXA, and to drink a cup of doubly labeled water containing harmless heavy water molecules. Individuals will be asked to collect small urine samples (about $^{1}/4$ cup) on the first day of the protocol and the last day of the protocol (10 d later). During this 10 d time period each subject will be asked to wear 3 accelerometers affixed to the hip, wrist and ankle.

4.a.2. Sources of Materials. Data for these studies will involve: 1) written survey data; 2) measurement of upper/lower extremity strength and function; 3) measurement of oxygen consumption; 3) measurement of movement data; and 4) anthropometric measures, and 5) measurement of energy expenditure via the double labeled water technique.

4.c.3. Potential Risks. Participants will be exposed to some risk during Study Aim 1. Survey measures pose minimal risk. Within Study Aim 1 there is risk associated with performing exercises and activities of daily living, but these will be self-selected. Risks include fatigue and muscle soreness. Some volunteers may experience claustrophobia in the room calorimeter. However, the room calorimeter at UC-AMC is large (12' x 12') with a large picture window, and we have had no reports of claustrophobia. To minimize the risk of participants experiencing claustrophobia, potential volunteers will be shown the calorimeter during visit 1 to familiarize them with the calorimeter.

Measurement of movement during Study Aim 2 poses no risk. The use of video recording represents the largest potential risk to consenting participants. Participant identity will be viewable, and because this takes place in natural settings there is the possibility that other individuals will also be in the recordings that are not part of the study. Precautions will be put in place, including that all direct observation video recordings will be immediately downloaded onto a password protected PC, and erased from the video recording device. Our customizable Noldus platform permits the coding team to blur any identities from the recording. This will be done for any non-participant as well as the participant themselves as identity is not important to the study outcomes. Extraneous persons captured from video that occurs in public settings is

permissible as there is no expectation of privacy. For those extraneous individuals captured in the home, private setting, sessions will only be recorded with their verbal consent, otherwise recording will be delayed until the participant is the only person in the frame in private settings. The recording device will only capture image, and no audio voice will be captured. Collectively these procedures have been approved by our institutional review boards for the protection of human subjects.

Risks associated with the doubly labeled water technique are minimal. All of the tracers used in this study are stable isotopes. There is no radiation hazard. Of these stable isotope tracers, only deuterium has a known biological toxicity. Behavioral effects are noted when 10-15% of body water is replaced with deuterium oxide and death occurs at 50% replacement. The highest level reached in this study is less than 0.1%. As with all research endeavors there is the potential risk of the loss of confidentiality.

4.b. Adequacy of Protection Against Risks

4.b.1. Recruitment and Informed Consent. Subjects within the designated age-range will be recruited from the metro Milwaukee area and the metro Denver area. All individuals who express an interest in participation will be screened for eligibility criteria. Informed consent (following institutional review) will be obtained from each participant, after thoroughly reviewing the procedures, risks and benefits of participation. All subjects will be informed that they may withdraw from the study at any time. After completing the informed consent process each participant will also be given a copy of the informed consent document for their personal records.

4.b.2. Protection Against Risks. While the risks associated with study participation are very minimal, the research team will put into place a number of precautions to address those minor risks:

- Collection of physiological data will take place at the CU-AMC and the UWM. Both sites are very familiar with all aspects of testing aforementioned in the Research Strategy. Extensive
- screening is in place to minimize the potential risks associated with research procedures.
- Familiarization time will be permitted with the 24-hr room calorimeter prior to study measurement.
- Video recording obtained from direct observation will be password protected. All faces on the recordings will be blurred using the Noldus Technology customizable platform. No recording will take place in a private setting that includes others in the frame without verbal consent. No audio will be collected.
- Only identification codes will be used to store and analyze data. Any identifying information will be kept in a locked secure site in a separate location to the housing of research data.
- All precautions will be taken to preserve the privacy and confidentiality of data.

4.c.3. Potential Benefits of the Proposed Research to the Subjects and Others. There are no direct benefits to the subjects who participate in Study Aims1, 2 or 3. However, the information gained from these studies has the potential to inform on more accurate representations of PA behavior in diverse populations. This information is particularly pertinent

to study the true associations between physical activity and health, in addition to being pertinent to clinical outcomes utilizing PA as a dependent variable.

4.3. Importance of Knowledge to be Gained. The majority of our understanding of the relationship between PA and health comes from studies employing PA assessment techniques in populations with no disability/functional limitations. There has been substantial development of PA assessment techniques in the general population with little attention given to those with movement limitations or movement impairments, even though 1 in 5 adults in the U.S. is thought to have a form of a disability. The hypothesis driven research in this proposal has tremendous potential to further our understanding of PA assessment in populations with disability/impairment.

INCLUSION OF WOMEN AND MINORITIES

5.a Inclusion of Women. Study enrollment and randomization will be stratified by gender and women will therefore represent 50% of our sample in all studies proposed.

5.b Inclusion of Minorities. We will be recruiting from an urban area, with large populations of ethnic minorities, including African American, Hispanic, American Indian, and Hmong. All recruitment advertisements will be sensitive to ethnic and racial differences. For instance, in posted announcements pictures or different ethnicities will be represented.

OMB Number: 0925-0002

PHS Inclusion Enrollment Report

Study Title: Calibrating free-living physical activity characteristics across functionally-limited populations using machine-learning accelerometer approaches

Racial Categories	Ethnic Categories									Total
	Not Hispanic or Latino			Hispanic or Latino			Unknown/ Not Reported Ethnicity			
	Female	Male	Unknown/ Not Reported	Female	Male	Unknown/ Not Reported	Female	Male	Unknown/ Not Reported	
American Indian/ Alaska Native	2	2		0	0					**4**
Asian	10	10		0	0					**20**
Native Hawaiian or Other Pacific Islander	0	0		0	0					**0**
Black or African American	22	21		4	4					**51**
White	120	116		30	29					**295**
More than One Race	5	5		0	0					**10**
Unknown or Not Reported										
Total	**159**	**154**		**34**	**33**					**380**

Study 1 of 1

INCLUSION OF CHILDREN

6.a Inclusion of Children. Children will be excluded from this study as they will not meet the age eligibility criteria of 18-100 years old.

RESOURCE SHARING PLAN

Data Management and Data Sharing Plan. In accordance with details within this proposal, data will be collected across two sites (UWM and UC-AMC) and data shared between these sites and UMass-Amherst. In conjunction with the timelines as proposed, considerable time will be dedicated to across site training using meticulous manuals of operation developed to maintain quality control and data similarities. All data will be kept on secured drives that are external to each laboratory and backed up in triplicate on a daily basis as part of the UWM and UC-AMC central research data servers and research drives.

In past projects, and recently completed projects, data has been securely shared through the sharing of encrypted and participant de-identified harddrives on a monthly basis. In the current proposal we plan to use **SPADES,** a multi-tier cloude infrastructure to store, visualize, share, and analyze data. This data storing and sharing resource has sufficient capacity for all generated data as outlined within this application. Please see accompanying letter of support from Fand Albinali.

Following completion of the Study Aims, all accrued data will be handled and collated in a manner consistent to allow for dissemination of generated data sets to qualified researchers in the scientific community for further analyses and study.