Where Are We Headed with Activity Monitors in Clinical Trials?

12th Annual
Patient-Reported Outcome Consortium Workshop

April 14-15, 2021

Disclaimer

- The views and opinions expressed in the following slides are those of the individual presenters and should not be attributed to their respective organizations/companies, the U.S. Food and Drug Administration or the Critical Path Institute.
- These slides are the intellectual property of the individual presenters and are protected under the copyright laws of the United States of America and other countries. Used by permission. All rights reserved. All trademarks are the property of their respective owners.

Session Participants

Moderator

- *Maria Mattera, MPH* – Assistant Director, PRO Consortium, C-Path

Presenters

- *Jennifer Goldsack, MChem, MA, MBA* – Executive Director, Digital Medicine Society (DiMe)
- *Bill Byrom, PhD* – Vice President, Product Intelligence and Positioning, Signant Health
- *Jiat Ling Poon, PhD* – Principal Research Scientist, Eli Lilly and Company
- *Milena Anatchkova, PhD* – Senior Research Leader, Evidera, Inc.

Additional Panelists

- *Andrew Potter, PhD* – Mathematical Statistician, Division of Biometrics I, Center for Drug Evaluation and Research, U.S. Food and Drug Administration
- *Steven Blum, MBA, MA* – Asset and Indication Lead, Patient-Reported Outcomes Assessment, WWHEOR, Bristol Myers Squibb

Session Outline

- Digital technology-based endpoints: Existing and emerging resources
- Implementation and endpoint considerations associated with activity monitors
- Case study: Highlights of the DIGItal Tracking of Arthritis Longitudinally (DIGITAL) Study
- Update from the Chronic Heart Failure (CHF) Working Group: Eliciting Meaningful Aspects of Physical Activity from Persons with CHF
- Panel Discussion
  - An FDA perspective
  - An industry perspective
- Question and Answer

DIGITAL MEDICINE SOCIETY

Digital technology-based endpoints: Existing and emerging resources

PRO Consortium Workshop

April 15, 2021

Jennifer Goldsack
Executive Director

12 Sponsors have collected digital endpoints

12 Sponsors have collected digital endpoints

Primary, Secondary or Label Claim

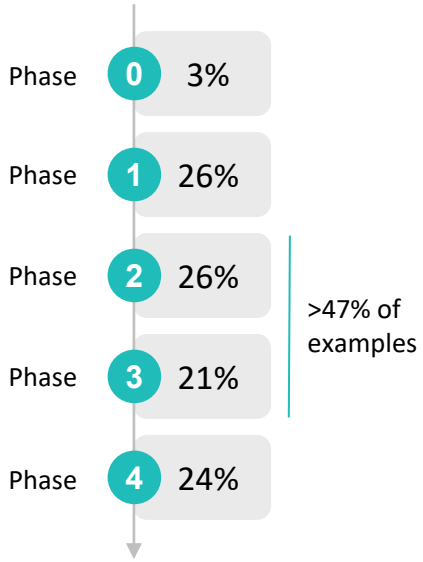
ACTELION, abbvie, Bellerophon THERAPEUTICS, Cavion, gsk, Pfizer, NOVARTIS, proteus DIGITAL HEALTH, REGENERON, SANOFI

Exploratory Only

Roche, teva

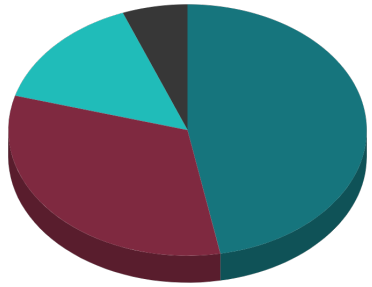
Sponsors start digital endpoint development early

Digital Endpoints



Digital endpoints are being used across drug, device, and biologic development

Investigational Product



Drug	47%
Device	32%
Biologic	15%
Other	6%

Pharma trusts digital products, primary/ secondary endpoints

Endpoint Positioning

16	Primary endpoints
4	Secondary endpoints
14	Exploratory
34	UNIQUE ENDPOINTS

DiMe DIGITAL MEDICINE SOCIETY

Is your company's work missing?
Submit it to DiMe:
https://bit.ly/DiMe-Endpoints

STAT FIRST OPINION

Digital endpoints library can aid clinical trials for new medicines

By JEN GOLDSACK, RACHEL A. CHASSE, *and* WILLIAM A. WOOD / NOVEMBER 6, 2019

52 Sponsors have collected digital endpoints

52 Sponsors have collected digital endpoints

Primary, Secondary or Label Claim

Abbott, abbvie, ABIDE THERAPEUTICS, ACTELION, ALKAHEST, ALTOIDA, AstraZeneca, AVAZIA, Bellerophon, BAYER, Cavion, dm, Daiichi-Sankyo, Eisai, gbt, GILEAD, gsk, janssen, Insulet, Hamad, Forest Laboratories, Inc., Genentech, LMC, MEDICIS, Lilly, MERCK, NANTCELL, NOVARTIS, novo nordisk, OLATEC, Pfizer, PrecisionBiotics, proteus, PURDUE, REGENERON, Sage Therapeutics, SANOFI, SENSIMED, SoniVie, Takeda, ucb, VANDA, VERTEX, ViiV Healthcare, ZOLL

Exploratory Only

CORTEXYME, Roche, sunovion, sonova HEAR THE WORLD, teva

Sponsors start digital endpoint development early

Digital Endpoints

Phase	0	1%
Phase	1	11%
Phase	2	34%
Phase	3	17%
Phase	4	37%

>51% of examples

Digital endpoints are being used across drug, device, and biologic development

Investigational Product

Drug	62.1%
Device	27.3%
Biologic	5.6%
Other	5%

Pharma trusts digital products, primary/ secondary endpoints

Endpoint Positioning

50	Primary endpoints
94	Secondary endpoints
22	Exploratory
166	UNIQUE ENDPOINTS

DiMe DIGITAL MEDICINE SOCIETY

Is your company's work missing?
Submit it to DiMe:
https://bit.ly/DiMe-Endpoints

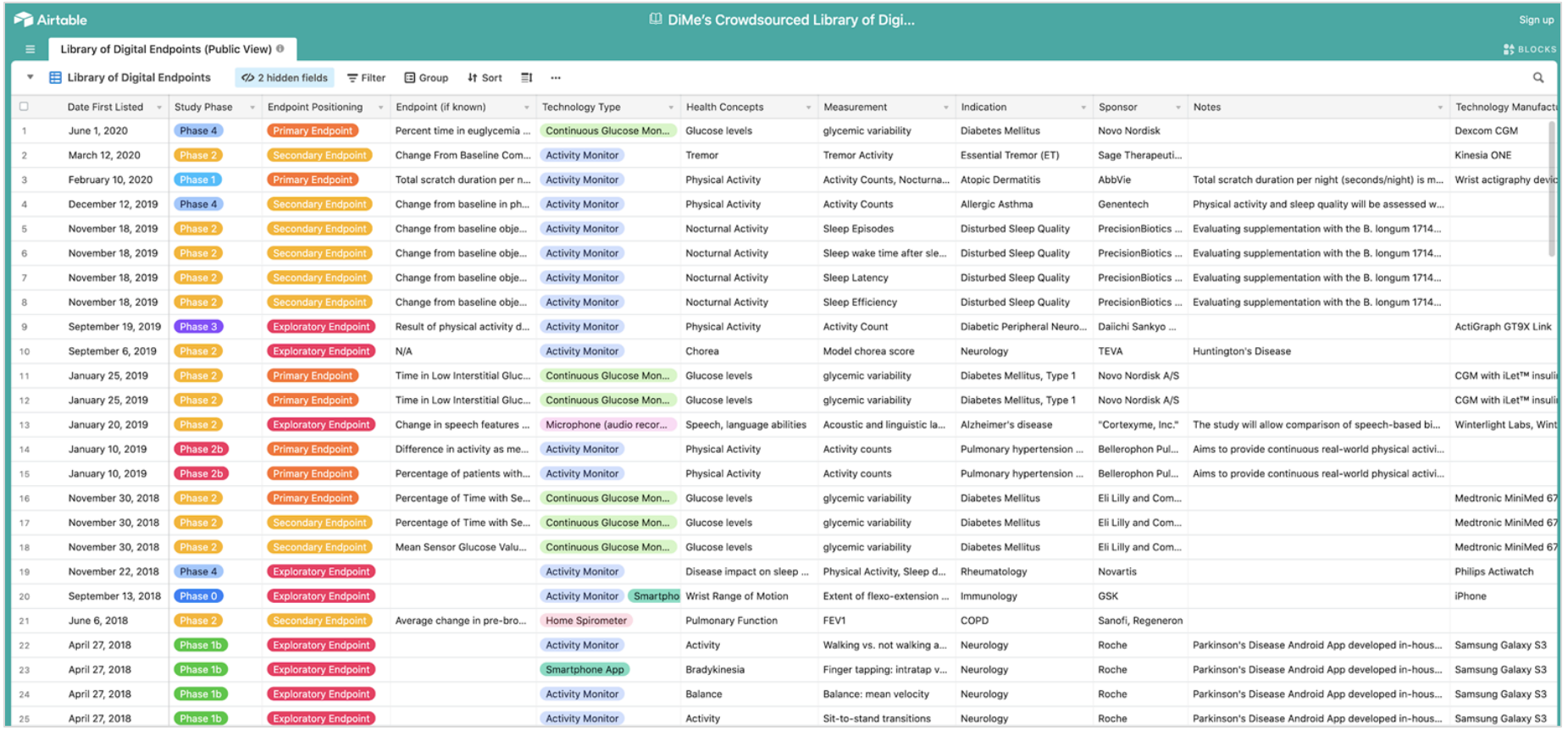
The Rapid Evolution of Digital Endpoints: Are We Headed in the Right Direction?

The number of unique digital endpoints being used in industry-sponsored trials of new medical products is skyrocketing, but is more always better?

Jennifer Goldsack

Jan 26 · 6 min read

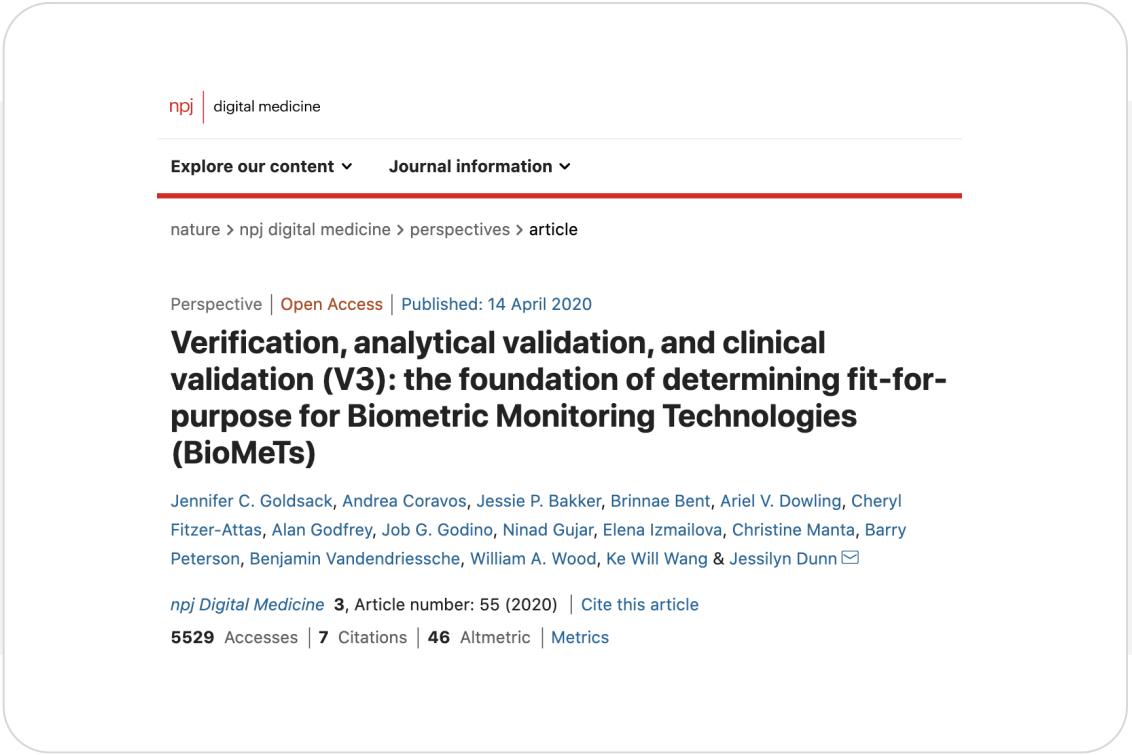
DiMe's Crowdsourced Library of Digital Endpoints



	Date First Listed	Study Phase	Endpoint Positioning	Endpoint (if known)	Technology Type	Health Concepts	Measurement	Indication	Sponsor	Notes	Technology Manufact
1	June 1, 2020	Phase 4	Primary Endpoint	Percent time in euglycemia ...	Continuous Glucose Mon...	Glucose levels	glycemic variability	Diabetes Mellitus	Novo Nordisk		Dexcom CGM
2	March 12, 2020	Phase 2	Secondary Endpoint	Change From Baseline Com...	Activity Monitor	Tremor	Tremor Activity	Essential Tremor (ET)	Sage Therapeuti...		Kinesia ONE
3	February 10, 2020	Phase 1	Primary Endpoint	Total scratch duration per n...	Activity Monitor	Physical Activity	Activity Counts, Nocturna...	Atopic Dermatitis	AbbVie	Total scratch duration per night (seconds/night) is m...	Wrist actigraphy devic
4	December 12, 2019	Phase 4	Secondary Endpoint	Change from baseline in ph...	Activity Monitor	Physical Activity	Activity Counts	Allergic Asthma	Genentech	Physical activity and sleep quality will be assessed w...	
5	November 18, 2019	Phase 2	Secondary Endpoint	Change from baseline obje...	Activity Monitor	Nocturnal Activity	Sleep Episodes	Disturbed Sleep Quality	PrecisionBiotics ...	Evaluating supplementation with the B. longum 1714...	
6	November 18, 2019	Phase 2	Secondary Endpoint	Change from baseline obje...	Activity Monitor	Nocturnal Activity	Sleep wake time after sle...	Disturbed Sleep Quality	PrecisionBiotics ...	Evaluating supplementation with the B. longum 1714...	
7	November 18, 2019	Phase 2	Secondary Endpoint	Change from baseline obje...	Activity Monitor	Nocturnal Activity	Sleep Latency	Disturbed Sleep Quality	PrecisionBiotics ...	Evaluating supplementation with the B. longum 1714...	
8	November 18, 2019	Phase 2	Secondary Endpoint	Change from baseline obje...	Activity Monitor	Nocturnal Activity	Sleep Efficiency	Disturbed Sleep Quality	PrecisionBiotics ...	Evaluating supplementation with the B. longum 1714...	
9	September 19, 2019	Phase 3	Exploratory Endpoint	Result of physical activity d...	Activity Monitor	Physical Activity	Activity Count	Diabetic Peripheral Neuro...	Daiichi Sankyo ...		ActiGraph GT9X Link
10	September 6, 2019	Phase 2	Exploratory Endpoint	N/A	Activity Monitor	Chorea	Model chorea score	Neurology	TEVA	Huntington's Disease	
11	January 25, 2019	Phase 2	Primary Endpoint	Time in Low Interstitial Gluc...	Continuous Glucose Mon...	Glucose levels	glycemic variability	Diabetes Mellitus, Type 1	Novo Nordisk A/S		CGM with iLet™ insuli
12	January 25, 2019	Phase 2	Primary Endpoint	Time in Low Interstitial Gluc...	Continuous Glucose Mon...	Glucose levels	glycemic variability	Diabetes Mellitus, Type 1	Novo Nordisk A/S		CGM with iLet™ insuli
13	January 20, 2019	Phase 2	Exploratory Endpoint	Change in speech features ...	Microphone (audio recor...	Speech, language abilities	Acoustic and linguistic la...	Alzheimer's disease	"Cortexyme, Inc."	The study will allow comparison of speech-based bi...	Winterlight Labs, Wint
14	January 10, 2019	Phase 2b	Primary Endpoint	Difference in activity as me...	Activity Monitor	Physical Activity	Activity counts	Pulmonary hypertension ...	Bellerophon Pul...	Aims to provide continuous real-world physical activi...	
15	January 10, 2019	Phase 2b	Primary Endpoint	Percentage of patients with...	Activity Monitor	Physical Activity	Activity counts	Pulmonary hypertension ...	Bellerophon Pul...	Aims to provide continuous real-world physical activi...	
16	November 30, 2018	Phase 2	Primary Endpoint	Percentage of Time with Se...	Continuous Glucose Mon...	Glucose levels	glycemic variability	Diabetes Mellitus	Eli Lilly and Com...		Medtronic MiniMed 67
17	November 30, 2018	Phase 2	Secondary Endpoint	Percentage of Time with Se...	Continuous Glucose Mon...	Glucose levels	glycemic variability	Diabetes Mellitus	Eli Lilly and Com...		Medtronic MiniMed 67
18	November 30, 2018	Phase 2	Secondary Endpoint	Mean Sensor Glucose Valu...	Continuous Glucose Mon...	Glucose levels	glycemic variability	Diabetes Mellitus	Eli Lilly and Com...		Medtronic MiniMed 67
19	November 22, 2018	Phase 4	Exploratory Endpoint		Activity Monitor	Disease impact on sleep ...	Physical Activity, Sleep d...	Rheumatology	Novartis		Philips Actiwatch
20	September 13, 2018	Phase 0	Exploratory Endpoint		Activity Monitor Smartpho	Wrist Range of Motion	Extent of flexo-extension ...	Immunology	GSK		iPhone
21	June 6, 2018	Phase 2	Secondary Endpoint	Average change in pre-bro...	Home Spirometer	Pulmonary Function	FEV1	COPD	Sanofi, Regeneron		
22	April 27, 2018	Phase 1b	Exploratory Endpoint		Activity Monitor	Activity	Walking vs. not walking a...	Neurology	Roche	Parkinson's Disease Android App developed in-hous...	Samsung Galaxy S3
23	April 27, 2018	Phase 1b	Exploratory Endpoint		Smartphone App	Bradykinesia	Finger tapping: intratap v...	Neurology	Roche	Parkinson's Disease Android App developed in-hous...	Samsung Galaxy S3
24	April 27, 2018	Phase 1b	Exploratory Endpoint		Activity Monitor	Balance	Balance: mean velocity	Neurology	Roche	Parkinson's Disease Android App developed in-hous...	Samsung Galaxy S3
25	April 27, 2018	Phase 1b	Exploratory Endpoint		Activity Monitor	Activity	Sit-to-stand transitions	Neurology	Roche	Parkinson's Disease Android App developed in-hous...	Samsung Galaxy S3

Which digital endpoint is right for your trial and how do you support it?

Evaluating digital clinical measures



npj | digital medicine

Explore our content ⌄ Journal information ⌄

nature › npj digital medicine › perspectives › article

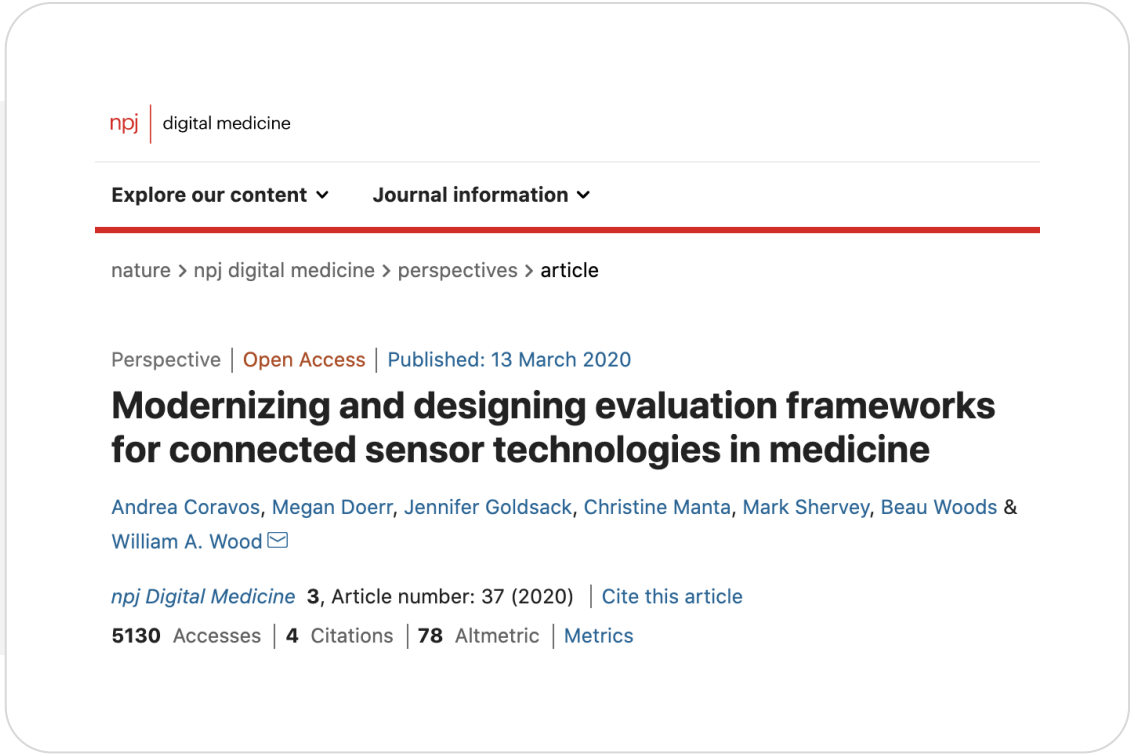
Perspective | Open Access | Published: 14 April 2020

Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs)

Jennifer C. Goldsack, Andrea Coravos, Jessie P. Bakker, Brinnae Bent, Ariel V. Dowling, Cheryl Fitzer-Attas, Alan Godfrey, Job G. Godino, Ninad Gujar, Elena Izmailova, Christine Manta, Barry Peterson, Benjamin Vandendriessche, William A. Wood, Ke Will Wang & Jessilyn Dunn ✉

npj Digital Medicine **3**, Article number: 55 (2020) | Cite this article

5529 Accesses | **7** Citations | **46** Altmetric | Metrics



npj | digital medicine

Explore our content ⌄ Journal information ⌄

nature › npj digital medicine › perspectives › article

Perspective | Open Access | Published: 13 March 2020

Modernizing and designing evaluation frameworks for connected sensor technologies in medicine

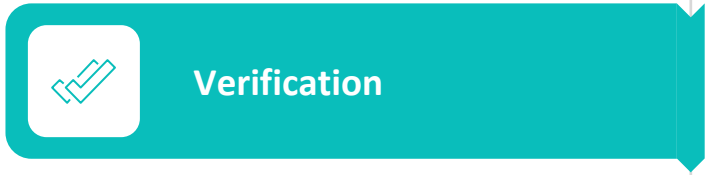
Andrea Coravos, Megan Doerr, Jennifer Goldsack, Christine Manta, Mark Shervey, Beau Woods & William A. Wood ✉

npj Digital Medicine **3**, Article number: 37 (2020) | Cite this article

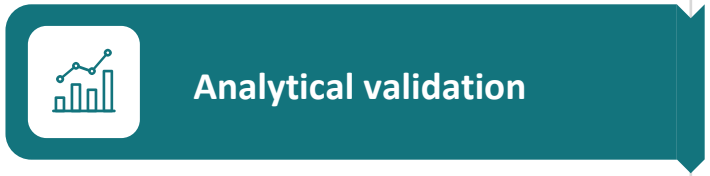
5130 Accesses | **4** Citations | **78** Altmetric | Metrics

V3 is a modular evaluation process

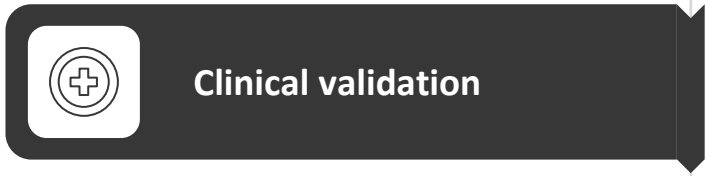
Design Specifications &
Modular Prototyping



Evaluates and demonstrates the performance of a sensor technology within a **BioMeT**, and the sample-level data it generates, against a pre-specified set of criteria.



Evaluates the performance of algorithm, and the ability of this component of the **BioMeT** to measure, detect, or predict physiological or behavioral metrics

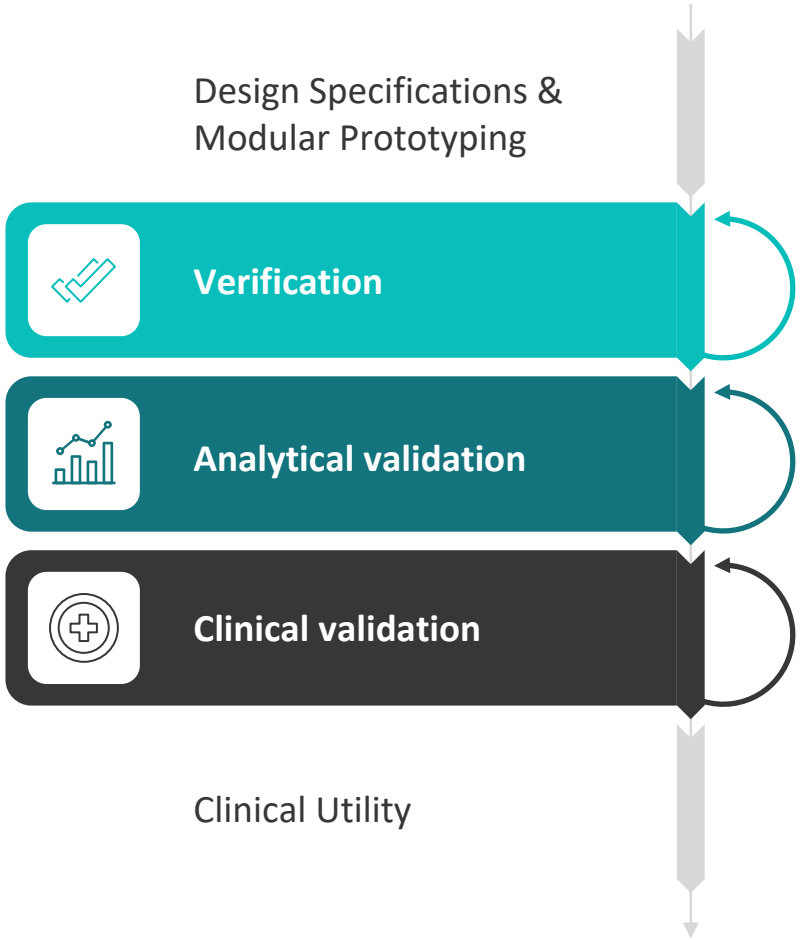


Evaluates whether a **BioMeT** acceptably identifies, measure, or predicts a meaningful clinical, biological, physical, functional state, or experience, in the stated context of use (which includes a specified population).

Clinical Utility

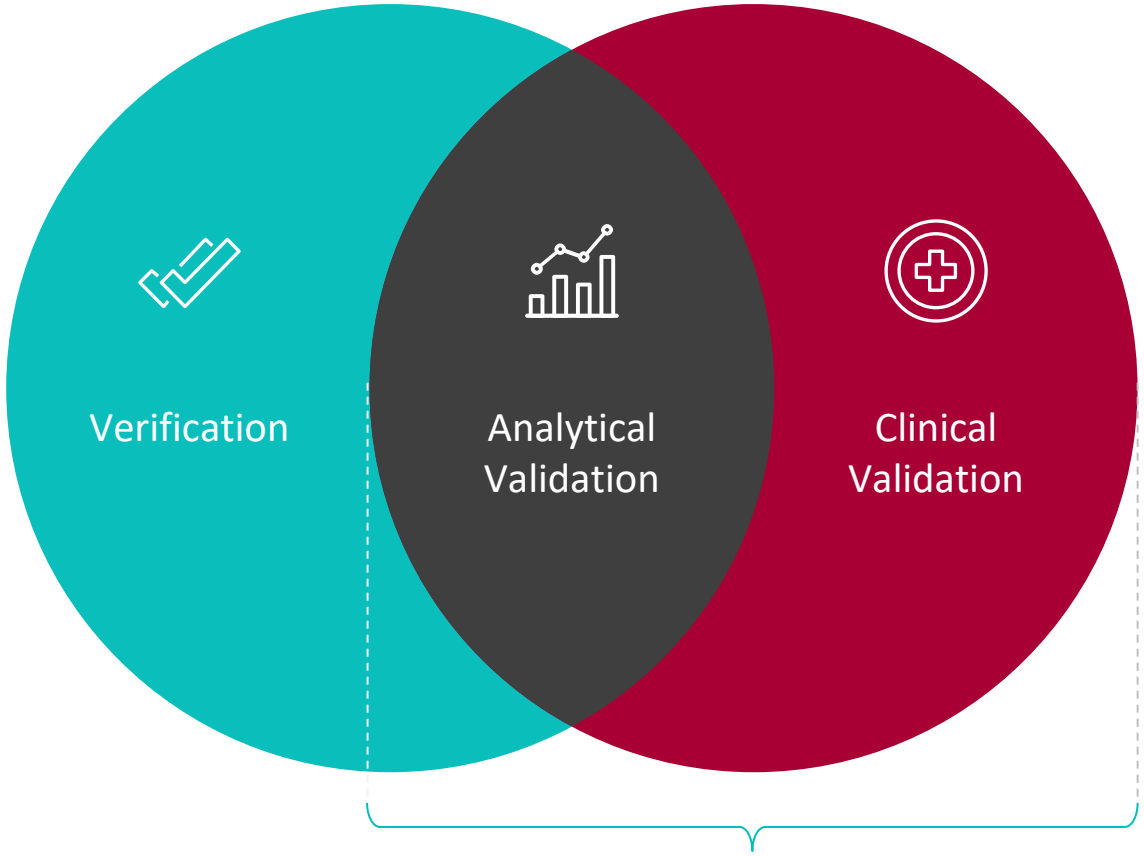
BioMeT - Biometric Monitoring Technology

Modular evaluation of digital measures



Changes to hardware/firmware?	• Reverification, or • Documentation of back-compatibility
Changes to software that change algorithm?	• Repeat analytical validation, or • Documentation of back-compatibility
Expansion to a new patient population?	• Repeat clinical validation if analytical validation in new population is documented, or • Repeat analytical & clinical validation

V3 processes are typically conducted by experts across disciplines and domains



Activity performed by:

(non-clinical) engineers

Both engineers and clinically-trained professionals

Clinically-trained professionals

Adoption of the V3 framework

nature reviews drug discovery

View all Nature Research journals

Explore our content ⌄ Journal information ⌄

nature › nature reviews drug discovery › comment › article

COMMENT · 29 SEPTEMBER 2020

Digital health technologies in clinical trials for central nervous system drugs: an EU regulatory perspective

Digital health technologies have the potential to help address some of the challenges in the clinical development of drugs for central nervous system disorders. This article discusses strategies for the development of such tools in the context of the European regulatory environment.

Valentina Mantua, Celso Arango, Pavel Balabanov & Florence Butlen-Ducuing ✉

Relating the V3 framework to current approaches

CTS Clinical and Translational Science
Open Access

Review | Open Access

Fit-for-Purpose Biometric Monitoring Technologies: Leveraging the Laboratory Biomarker Experience

Alan Godfrey, Benjamin Vandendriessche, Jessie P. Bakker, Cheryl Fitzer-Attas, Ninad Gujar, Matthew Hobbs, Qi Liu, Carrie A. Northcott, Virginia Parks, William A. Wood, Vadim Zipunnikov, John A. Wagner, Elena S. Izmailova ... See fewer authors

First published: 08 August 2020 | https://doi.org/10.1111/cts.12865

Relating the V3 framework to current approaches

Review > Contemp Clin Trials. 2020 Apr;91:105962. doi: 10.1016/j.cct.2020.105962.
Epub 2020 Feb 20.

Considerations for development of an evidence dossier to support the use of mobile sensor technology for clinical outcome assessments in clinical trials

M K Walton [1], J C Cappelleri [2], B Byrom [3], J C Goldsack [4], S Eremenco [5], D Harris [6], E Potero [7], N Patel [8], E Flood [8], M Daumer [9]

V3 is the first step of a comprehensive evaluation framework for fit-for-purpose connected sensors

Verification, Analytical Validation and Clinical Validation (V3)	Does the tool measure what it claims to measure? Is the measurement appropriate for the target population?
Security	Does the manufacturer build with safety by design? Is there a Disclosure Policy? Software Bill of Materials?
Data Rights and Governance	Who has access to the data and when? Is the privacy policy publicly accessible?
Utility and Usability	How is the tool worn? Battery life? Available technical support?
Economic Feasibility	What's the net benefit versus price? Is cost a one-time or subscription model?

And none of this matters if you didn't pick a measure that matters in the first place

Digital Biomarkers

Digit Biomark 2020;4:69–77

DOI: 10.1159/000509725
Received: May 9, 2020
Accepted: June 25, 2020
Published online: September 15, 2020

© 2020 The Author(s)
Published by S. Karger AG, Basel
www.karger.com/dib

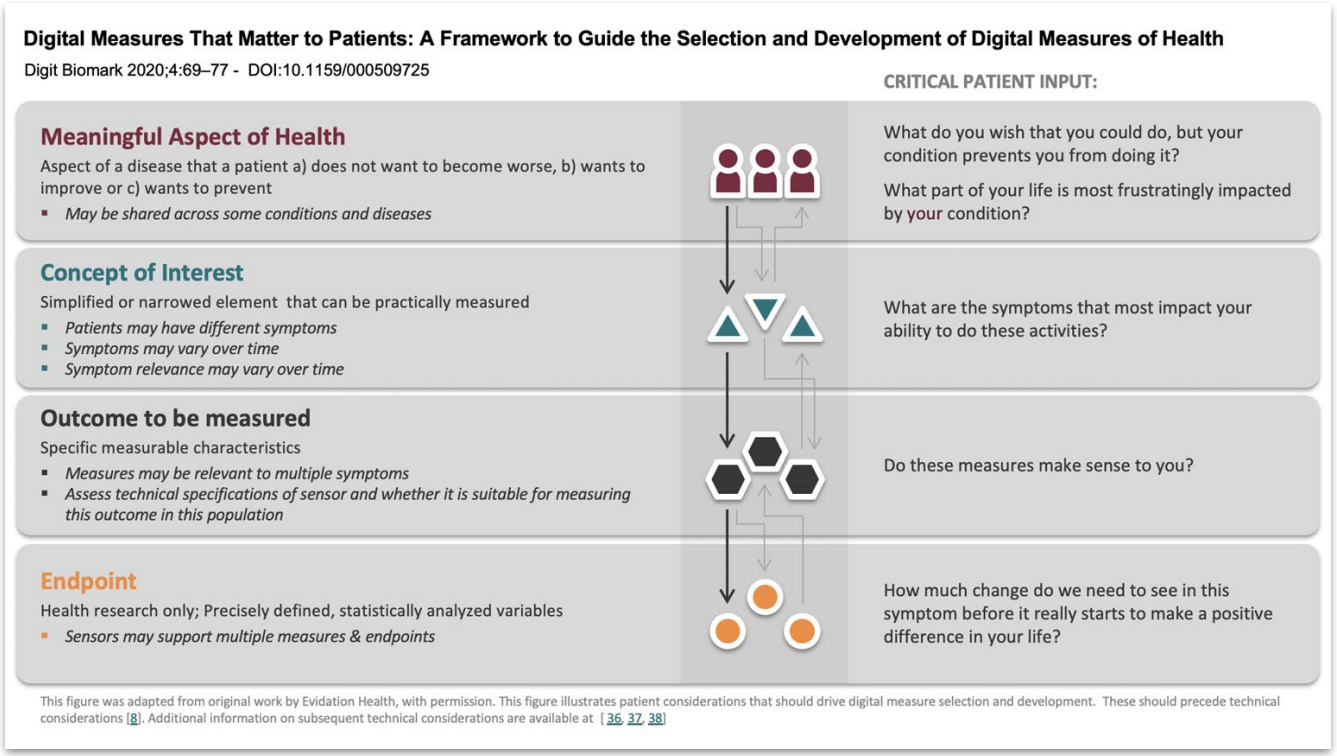
This article is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND) (http://www.karger.com/Services/OpenAccessLicense). Usage and distribution for commercial purposes as well as any distribution of modified material requires written permission.

Viewpoint – Review Article

Digital Measures That Matter to Patients: A Framework to Guide the Selection and Development of Digital Measures of Health

Christine Manta[a, b] Bray Patrick-Lake[a, c] Jennifer C. Goldsack[a]

[a]Digital Medicine Society, Boston, MA, USA; [b]Elektra Labs, Boston, MA, USA; [c]Evidation Health, Inc., San Mateo, CA, USA



Source: https://www.karger.com/Article/Pdf/509725

TOUR OF DUTY: DRIVING ADOPTION

The Playbook: Digital Clinical Measures

Terminology supporting the essential guide for successful remote monitoring across *clinical research*, *clinical care*, and *public health*.

ActiGraph · Activ insights · APhA · Big Ideas Lab · BlackThorn Therapeutics · Center for Digital Health · COVANCE Solutions Made Real · DiGITAL.HEALTH · DiME Digital Medicine Society · elektralabs · Lilly · European Medicines Agency · evidation · FDA · Genentech A Member of the Roche Group · koneksa · MERCK · MYOKARDIA · Open mHealth · Pfizer · physIQ · savvy · SageBionetworks · Scripps Research · Takeda · VivoSense · WINTERLIGHT

The Playbook will walk you through the foundational processes for successful remote monitoring

Measures

1. Determine the **meaningful aspect of health** (MAH)
2. Identify the **concept of interest** (COI)
3. Define the **digital measure** (e.g, outcome/endpoint)

Technologies

Evaluate the **risk/benefit** to ensure safety and efficacy (e.g., validation (V3), utility & usability, security, data rights)

Operations

Plan for the **jobs to be done** during deployment (e.g., purchasing, distribution, monitoring, data analysis)

Opportunities for **collaboration across industry** include**:**

- Promoting a culture of **ethics** to ensure equality and justice
- Developing **benchmarks** to compare digital measures (e.g., algorithms)
- Setting and developing **standards** for digital measures
- Participating in the **policy and regulatory** process (e.g. public comments)

Source: Playbook team analysis

Following *The Playbook* flow, you'll avoid order-related pain points

OPTIMAL FLOW

Measures

"Parkinson's patients tell us that being able to walk independently is important, so we're interested in measuring activity"

Technologies

"Using a smartwatch with accelerometer"

Operations

"Deploying 250 smartwatches across 15 countries"

COMMON ISSUE #1

Technologies

"I saw a cool Apple watch at a Conference"

Operations

"I ordered 250 of them! Shipping to our patients."

Measures

2 years later reviewing results: "Discard data. No clinically relevant signals yet in Parkinson's"

COMMON ISSUE #2

Operations

"We're deploying 250 smartwatches across 15 countries"

Technologies

"Patients are having Bluetooth connection issues; security incident. Compliance is down"

Measures

"We weren't able to gather any usable data to construct a measure"

Source: Playbook team analysis

PRO TIP

Regulatory approval of a technology does **not** necessarily indicate *fit-for-purpose*

FDA clearance of a technology and/or the presence of a CE mark should not be used in place of the evaluation processes described in *The Playbook* to determine the suitability of a technology for use in remote monitoring during a clinical trial.

Source: Playbook team analysis

In the U.S., you do *not* need to use an FDA-cleared product to collect a digital measurement in a registered clinical trial

Does a 'device' need to have an FDA 510(k), De Novo, or PMA to be used in a clinical trial?

No, this is one of the most common misperceptions.

Mobile technologies for data capture in clinical trials do not necessarily need to be approved or cleared as a medical device.

If the technology manufacturer does not make a medical claim, it is not regulated as a device.

SPOTLIGHT

- **FDA Public Workshop on Endpoints for Regulatory Decision Making (Dec '19):** See more in this video at timestamps 31:25, 38:48 and 55:54
- **FiercePharma Webinar** with FDA Officials on 510k not required**:** Watch here
- **Clinical Trials Transformation Initiative (CTTI)** a FDA-Duke, public-private partnership on Mobile Technologies (see page 3)
- **Digital Medicine Society (DiMe)** has posted a public comment to ask the FDA to formally express clarify this misperception in a guidance

Source: Playbook team analysis

In Europe, you do *not* need to use an CE-marked product to collect a digital measurement in a registered clinical trial

A commercial technology likely does not require a CE mark if it is:

1. Being used as a **research tool** in drug development, **and**
2. **Not intended for commercialization** associated with any medical claim

Drug developers should to discuss their plans to use digital measurement product with the EMA, and specifically any aspects that may impact the benefit risk evaluation of a new drug by EMA.

A reminder on the European Medicine Agency's (EMA) remit: the agency only has jurisdiction over the benefit-risk of a medicine, but not devices.

SPOTLIGHT

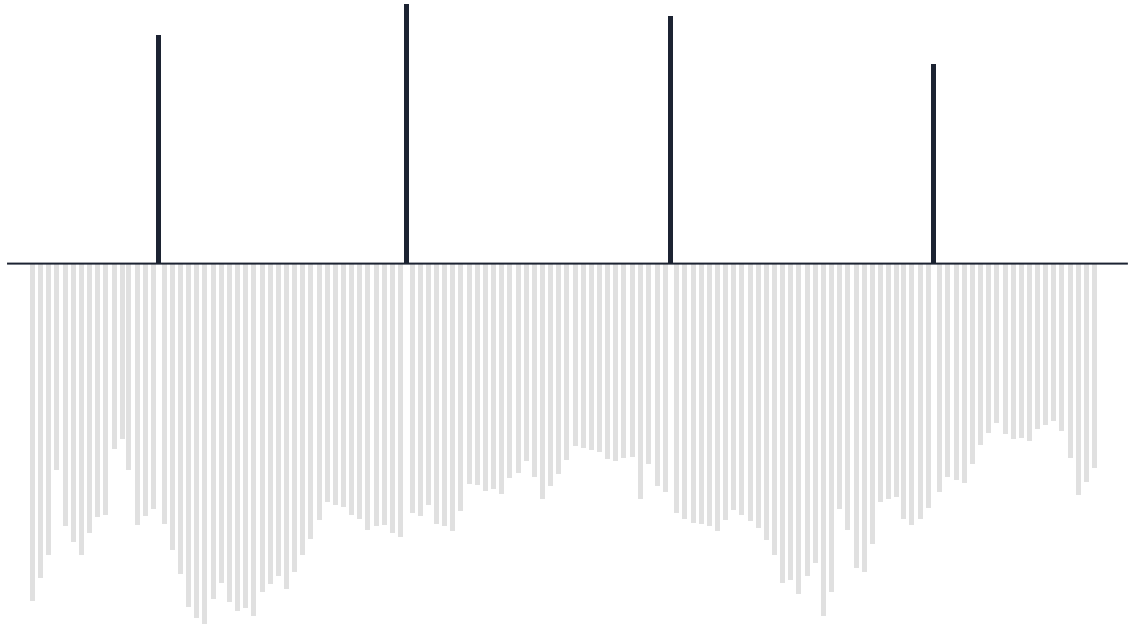
European Medicines Agency. "Questions and answers: Qualification of digital technology-based methodologies to support approval of medicinal products." (2020)

Cerreta, Francesca, et al. "Digital technologies for medicines: shaping a framework for success." (2020)

Source: Embedded above, Playbook team analysis

Remote monitoring using connected sensors offers *a more holistic view* of a person's lived experience

Visible Data Points
(episodic)



Invisible Data Points
(continuous)

Data collected from **traditional** visits to hospitals and clinics

Data collected **remotely** from everyday life

Source: "Visible vs. Invisible Data" chart designed by Evidation Health, re-worked by Elektra Labs, Playbook team analysis | playbook.dimesociety.org

DIGITAL MEDICINE SOCIETY

THANK YOU

Jennifer Goldsack

www.dimesociety.org

@_DiMeSociety

linkedin.com/company/dime-society

Where Are We Headed with Activity Monitors in Clinical Trials?

Implementation and endpoint considerations

Bill Byrom, PhD
VP Product Intelligence and Positioning
Signant Health

Challenges and gaps in knowledge

Are we measuring something meaningful?

Do we have good implementation and data management standards?

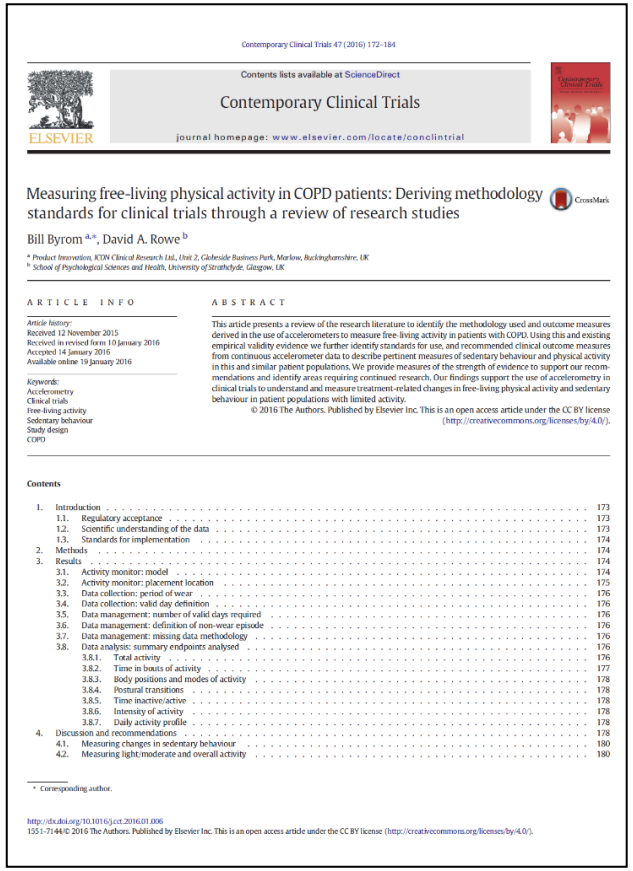
Do we understand the relationship between endpoint properties and the specific device?

Do we understand the evidence needed to support device selection?

Implementation and data management considerations

Challenge determining consensus standards: an example

- Review of implementation methodology and outcomes measures
- 76 COPD studies
- Many different placement locations
  - Waist/hip (41%), arm (15%), ankle (8%), wrist (4%), lower back (3%), pocket (1%), shoe (1%), multiple sensors (9%), not reported (18%)
- Different periods of wear studied
  - 2 days to 26 weeks (median 7 days)
- Data management rules
  - Valid day: 6 – 24 hours (10-12 h most common)
  - Number of valid days needed: most common – 3 to 7 days
  - Non-wear detection rule: most common – 60 minutes zero counts
- > 80 endpoints reported



Contemporary Clinical Trials 47 (2016) 172–184

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Measuring free-living physical activity in COPD patients: Deriving methodology standards for clinical trials through a review of research studies

Bill Byrom [a,*], David A. Rowe [b]

[a] Product Innovation, ICON Clinical Research Ltd., Unit 2, Globeside Business Park, Marlow, Buckinghamshire, UK
[b] School of Psychological Sciences and Health, University of Strathclyde, Glasgow, UK

ARTICLE INFO

Article history:
Received 12 November 2015
Received in revised form 10 January 2016
Accepted 14 January 2016
Available online 19 January 2016

Keywords:
Accelerometry
Clinical trials
Free-living activity
Sedentary behaviour
Study design
COPD

ABSTRACT

This article presents a review of the research literature to identify the methodology used and outcome measures derived in the use of accelerometers to measure free-living activity in patients with COPD. Using this and existing empirical validity evidence we further identify standards for use, and recommended clinical outcome measures from continuous accelerometer data to describe pertinent measures of sedentary behaviour and physical activity in this and similar patient populations. We provide measures of the strength of evidence to support our recommendations and identify areas requiring continued research. Our findings support the use of accelerometry in clinical trials to understand and measure treatment-related changes in free-living physical activity and sedentary behaviour in patient populations with limited activity.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

1. Introduction . . . 173
1.1. Regulatory acceptance . . . 173
1.2. Scientific understanding of the data . . . 173
1.3. Standards for implementation . . . 174
2. Methods . . . 174
3. Results . . . 174
3.1. Activity monitor: model . . . 174
3.2. Activity monitor: placement location . . . 175
3.3. Data collection: period of wear . . . 176
3.4. Data collection: valid day definition . . . 176
3.5. Data management: number of valid days required . . . 176
3.6. Data management: definition of non-wear episode . . . 176
3.7. Data management: missing data methodology . . . 176
3.8. Data analysis: summary endpoints analysed . . . 176
3.8.1. Total activity . . . 176
3.8.2. Time in bouts of activity . . . 177
3.8.3. Body positions and modes of activity . . . 178
3.8.4. Postural transitions . . . 178
3.8.5. Time inactive/active . . . 178
3.8.6. Intensity of activity . . . 178
3.8.7. Daily activity profile . . . 178
4. Discussion and recommendations . . . 178
4.1. Measuring changes in sedentary behaviour . . . 180
4.2. Measuring light/moderate and overall activity . . . 180

* Corresponding author.

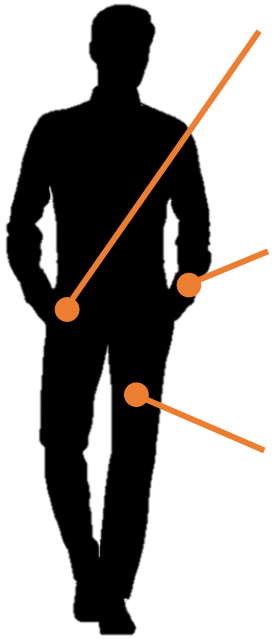
http://dx.doi.org/10.1016/j.cct.2016.01.006
1551-7144/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contemporary Clinical Trials, 2016; 47:172-184.

COPD: chronic obstructive pulmonary disease

Things to consider

PLACEMENT LOCATION



HIP
- Close to centre of mass
- Good specificity for SED, LPA, MVPA
- Best for energy expenditure in lab setting

WRIST
- Poorer at distinguishing LPA
- More PA across all intensities, less SED
- Better wear compliance

THIGH
- Less attenuation of ground reaction force
- Inclinometer
  - Superior for SED
  - Able to distinguish walking vs walking upstairs

WEAR TIME

ENDPOINT-SPECIFIC

DAILY TOTAL
- Total Steps
- Total MVPA time
- Measure for a good proportion of the day to get a reliable estimate (e.g., 10-14 h)
- Measure for a reasonable number of days (e.g., 5+ days)
  - Number of weekend days?

SPEED/CADENCE
- Consistent across bouts
- Measure one of more suitable bouts over the day

SAMPLE RATE

HZ
- 60 – 100 Hz

EPOCH LENGTH
- 15 s for SED
- 15 – 60 s for PA

DATA MANAGEMENT RULES

- Definition of a valid day
- Number of valid days needed
- Rules for missingness
- Detection of non-wear periods
  - Children vs adults

SED: sedentary behaviour
PA: physical activity
LPA: light physical activity
MVPA: moderate to vigorous physical activity

Are endpoints device-agnostic?

Important questions

- Are measurement properties sensor-specific?
  - Can I interchange sensors?
  - Can I compare / pool the results from multiple studies?
    - Same endpoint, different device
    - Can I interpret changes on a common scale?
  - What about “bring-your-own-wearable” (BYOW)?

Actual state: consumer devices

This full text paper was peer-reviewed at the direction of IEEE Instrumentation and Measurement Society prior to the acceptance and publication.

Measuring the Fitness of Fitness Trackers

Chelsea G. Bender, Jason C. Hoffstot
Brian T. Combs and Sara Hooshangi
Integrated Information, Science, and Technology
The George Washington University
Washington, DC, USA
Email: shoosh@gwu.edu

Justin Cappos
Computer Science and Engineering
New York University
New York, NY, USA
Email: jcappos@nyu.edu

***Abstract*—Data collected by fitness trackers could play an important role in improving the health and well-being of the individuals who wear them. Many insurance companies even offer monetary rewards to participants who meet certain steps or calorie goals. However, in order for it to be useful, the collected data must be accurate and also reflect real-world performance. While previous studies have compared step counts data in controlled laboratory environments for limited periods of time, few studies have been done to measure performance over longer periods of time, while the subject does real-world activities. There are also few direct comparisons of a range of health indicators on different fitness tracking devices. In this study, we compared step counts, calories burned, and miles travelled data collected by three pairs of fitness trackers over a 14-day time period in free-living conditions. Our work indicates that the number of steps reported by different devices worn simultaneously could vary as much as 26%. At the same time, the variations seen in distance travelled, based on the step count, followed the same trends. Little correlation was found between the number of calories burned and the variations seen in the step count across multiple devices. Our results demonstrate that the reporting of health indicators, such as calories burned and miles travelled, are heavily dependent on the device itself, as well as the manufacturer's proprietary algorithm to calculate or infer such data. As a result, it is difficult to use such measurements as an accurate predictor of health outcomes, or to develop a consistent criteria to rate the performance of such devices in head-to-head comparisons.**

Keywords—Fitness Trackers; Accuracy; Physical Activity; Free-living Conditions

I. Introduction

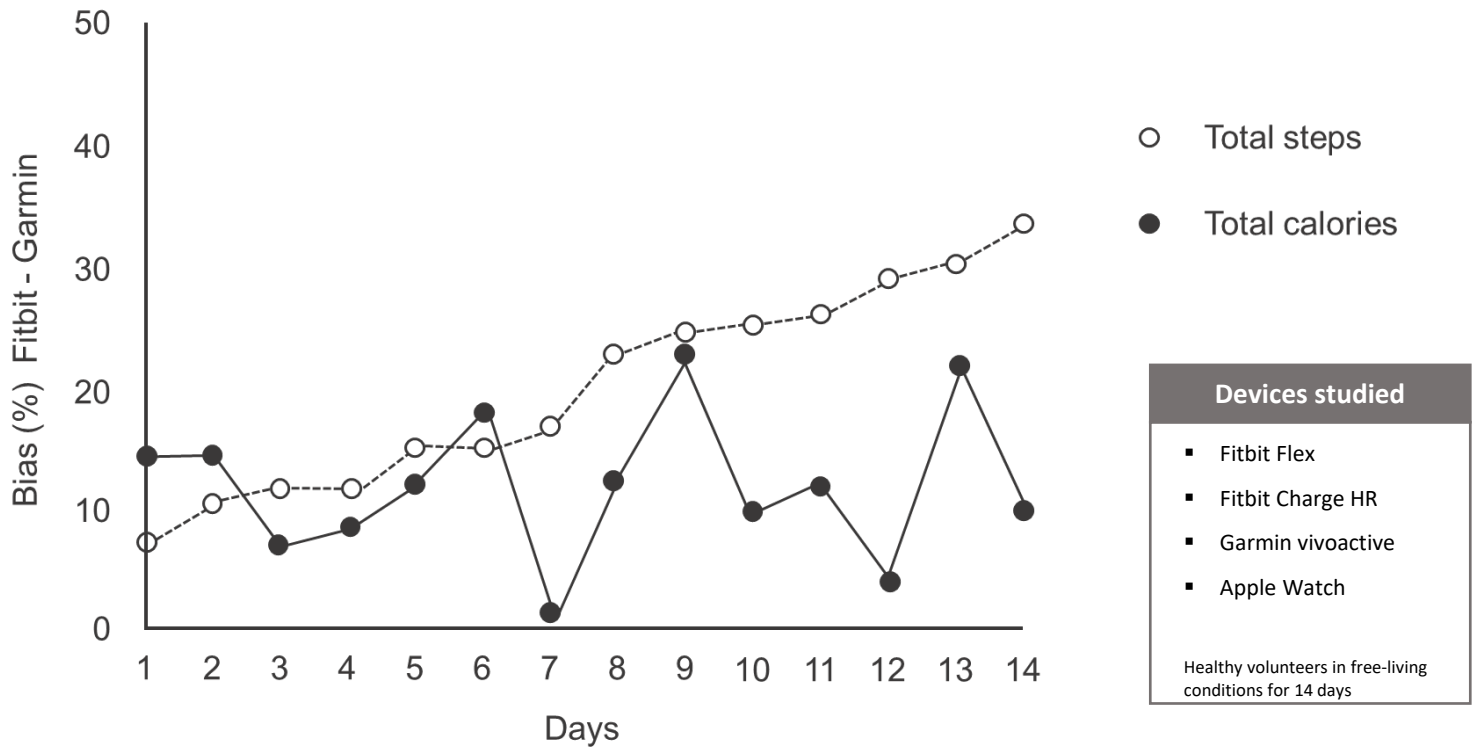
The past several years have seen an exponential growth in the market for personal wearable devices, with estimated sales of up to 126 million units anticipated by the end of 2019 [1]. Fitness tracking devices lead sales in this market, and continue to gain popularity as the correlation between an active lifestyle and the prevention of chronic diseases is demonstrated by research [2], [3]. These trackers give their users the ability to monitor and track key health markers, thus encouraging them to continue their healthy efforts.

As manufacturers try to improve the accuracy of these health measurements by adding functionality and introducing new devices into the market place at a rapid pace, the number of ways that this collected and stored data can be used also increases. Individuals can use data on their average daily/weekly physical activity to monitor their own health, or to identify key markers to report to their health providers. Public health researchers could use such data in aggregated form, in large-scale studies to monitor health related outcomes for different segments of the population. And, on a larger scale, programs sponsored by insurance companies can promote healthier lifestyles by offering incentivizing discounts on life and health insurance products based on the physical activity levels of consumers.

Such programs, however, rely on the ability of these devices to reliably generate accurate data. Data accuracy ultimately depends on two factors: the quality of the sensors embedded in the device, and the algorithm used to interpret the raw data. To this end, there has been a surge in the number of research studies testing the accuracy of wearable fitness devices as compared to research-grade accelerometers and multi-sensor devices [4]–[11]. Most of these studies have focused on a cross-sectional comparison of consumer-based products to research-grade gold standards only in a laboratory or a controlled real-world environment [4]–[7], [12]. Conducting experiments without the prescribed restrictions of a laboratory (i.e. under a free-living condition) is significantly more challenging, as the variations in speed, direction and intensity of physical activities are larger. This may be why only a few studies have measured the accuracy of trackers in free-living conditions [8]–[10] and most free-living studies have been short in duration, typically in the range of one or two days. Furthermore, the integrity of these results could also be compromised if the subjects under study (who are often volunteers) do not follow the experiment protocols.

In this work, we set out to compare parameters and experimental settings that have not been explored in previous work. We start by looking at other health indicators measured by these devices, such as calories burned or distance travelled. We designed a series of experiments to compare these along with the more commonly studied measure of step counts. While the step count provides a general sense of movement and physical activity, calories burned and the number of miles travelled could be better indicators of an individual's energy expenditure and, hence his/her physical fitness level. If the fitness trackers are to become an integral part of our health-

Bender CG, Hoffstot JC, Combs BT et al. Measuring the Fitness of Fitness Trackers. Sensors Applications Symposium 2017 IEEE, pp. 1-6, 2017, March.



Devices studied

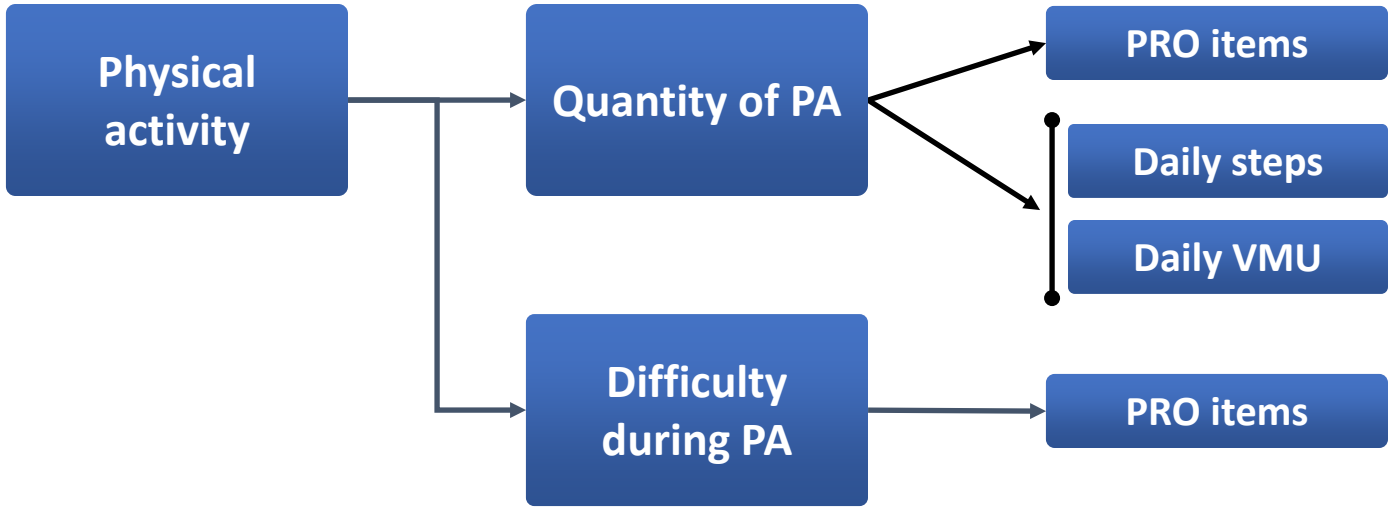
- Fitbit Flex
- Fitbit Charge HR
- Garmin vivoactive
- Apple Watch

Healthy volunteers in free-living conditions for 14 days

> Step count, distance travelled, and calories burned could vary significantly between devices used concurrently.

Actual state: research grade devices

- PROActive measure
  - Clinical visit PROactive Physical Activity in COPD (C-PPAC) tool
  - Daily PROactive Physical Activity in COPD (D-PPAC) tool



McRoberts Dynaport

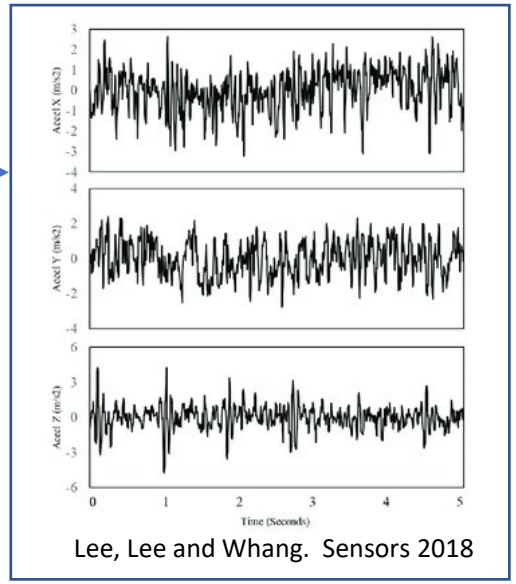
ActiGraph GT9X

Composite endpoint:
Monitor-specific data conversion

PA: physical activity
VMU: vector magnitude unit (a measure of PA intensity)

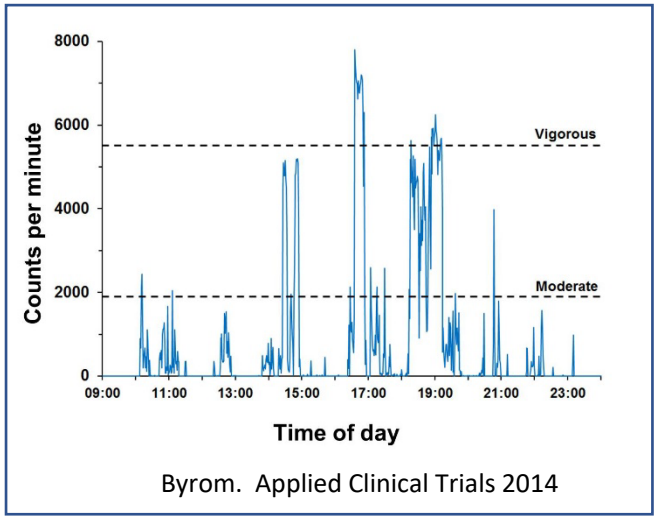
Types of data

1. Raw signal
   - The raw accelerations measured using the tri-axial sensor.



Lee, Lee and Whang. Sensors 2018

2. Epoch level data
   - Summarised raw acceleration data into epoch level (e.g., 1 minute) time series, often known as activity counts.



Byrom. Applied Clinical Trials 2014

3. Daily summaries
   - Summary outcome measures such as total number of steps, time spent in bouts of activity / sedentary behaviour etc.

Fitbit:
- Total calories
- Total steps
- Total floors
- Total distance
- Minutes: sedentary, light, fairly active, very active

The potential for device-agnostic endpoints

1. Raw signal
   - The raw accelerations measured using the tri-axial sensor.

2. Epoch level data
   - Summarised raw acceleration data into epoch level (e.g., 1 minute) time series, often known as activity counts.

3. Daily summaries
   - Summary outcome measures such as total number of steps, time spent in bouts of activity / sedentary behaviour etc.

Use raw data and apply a common processing routine

- Minor differences in how manufacturers process the data at device level prior to transfer, but much less than epoch level and summary data
- Challenge of harmonising the data in terms of structure, etc.
- Raw data rarely available
  - Usually available for research grade devices e.g., Actigraph, Axivity, ActiPal, Verisense, and others.
  - Rarely available from consumer devices such as Fitbit, Garmin or Withings
    - Some collaborations have been possible, but not available through standard APIs/SDKs.

Understand the tolerances in outcome measures and how they affect the endpoint measure

API: application programming interface
SDK: software development kit

Thank you

@billbyrom bill.byrom@signanthealth.com

PRO CONSORTIUM

Highlights of the DIGItal Tracking of Arthritis Longitudinally (DIGITAL) Study

Jiat Ling Poon, PhD
Principal Research Scientist
Eli Lilly and Company

Study Objectives

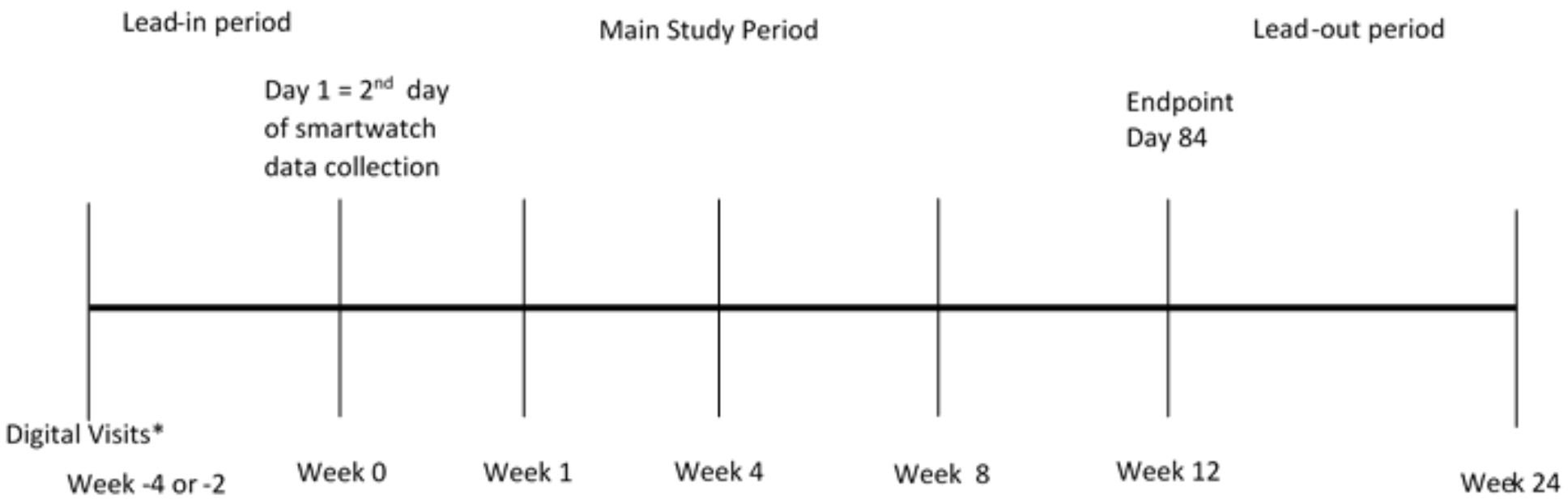
- **Primary:** Assess agreement between digital sensor measures from Fitbit® Versa™ (smartwatch) and PRO measures from participants with rheumatoid arthritis (RA)
- **Secondary:**
  - Explore factors associated with patient-reported worsening
  - Evaluate reliability and validity of smartwatch measures
  - Explore factors associated with adherence to technology
  - Evaluate and outline operational processes

Nowell WB et al. (2019) JMIR Res Protoc;8(9):e14665

Methods: Study Population

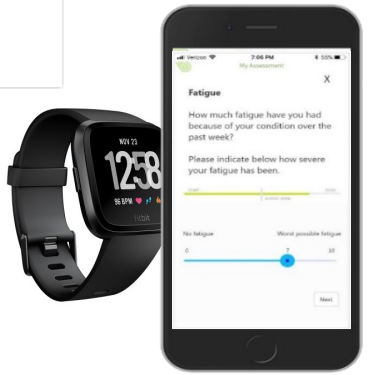
- Key inclusion criteria:
  - Age 19 or older
  - U.S. resident
  - Self-reported diagnosis of RA
  - Registered a valid e-mail address with ArthritisPower
  - Currently being seen by a U.S. rheumatologist
  - Currently taking at least one conventional synthetic or targeted disease-modifying antirheumatic drug for RA, but not baricitinib
  - Already own a smartphone (iPhone or Android) to which they have downloaded the ArthritisPower app

Methods: Study Schema



* Active data collection of 2 daily questionnaires begins 2-4 weeks prior to receipt of smartwatch and continues throughout the study. Other ePROs will be collected weekly. Digital data from the smartwatch will be collected passively.

ArthritisPower®



Methods: Data Collection and Analysis

Patient-reported Outcome Measures

- Weekly:
  - OMERACT RA flare questionnaire
  - PROMIS CAT:
    - Fatigue
    - Pain interference
    - Physical function
    - Satisfaction with participation in discretionary social activities
    - Sleep disturbance
  - Godin Leisure-Time Physical Activity Questionnaire
- Daily:
  - Pain NRS
  - Fatigue NRS

Selected smartwatch measures

- Activity
  - Steps
  - Activity intensity
- Heart rate
  - Beats per minute
  - Time in heart rate zone of interest
- Sleep
  - REM
  - Sleep efficiency
  - Awake time
  - Awakenings

Analysis for primary objective: Pearson's correlation coefficient to assess agreement between PRO and smartwatch measures

NRS: numeric rating scale; OMERACT: Outcome Measures in Rheumatology; PROMIS CAT: Patient-Reported Outcomes Measurement Information System computer adaptive tests; RA: rheumatoid arthritis; REM: rapid eye movement

Pilot (soft launch) Participant Feedback

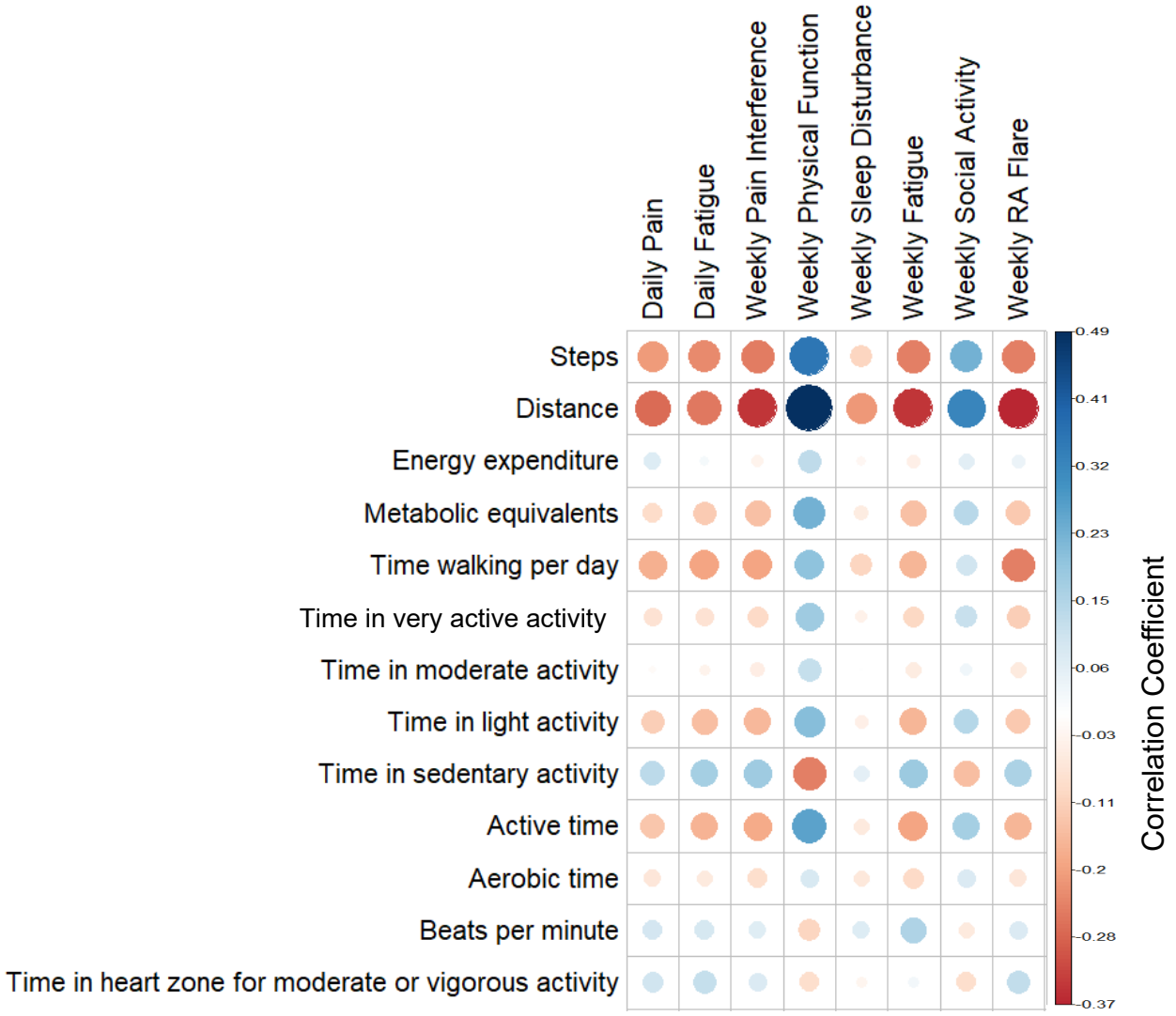
- **Assessments and tracking**
  - Participants liked being able to see their own (limited) data to understand their sleep and activity patterns
  - Participants liked being able to see and track their results over time to be more self-aware of their activities, symptom fluctuations and/or improvements, relationship between various symptoms they experienced (e.g., between pain and fatigue)
  - Some participants experienced confusion and difficulties over timing the completion of their daily assessments if they were not entering data at the same time each day
- **Smartwatch setup and use**
  - Initial study setup was confusing for some participants due to unclear instructions
  - Participants who were not used to charging their watch or wearing a watch at all times sometimes forgot to put their watch back on after charging, or did not wear their watch in the shower/while sleeping
  - Some participants experienced discomfort wearing their watch at all times – sensitive skin, moisture trapped under the watch close to the skin, etc.
- **Notifications/reminders**
  - Most participants appreciated the daily/weekly reminders to complete their study assessments or sync their devices, with some suggesting that more frequent reminders would be helpful

DIGITAL Study Participant Demographic and Clinical Characteristics

Qualified for Main Study Period (N=278)	n (%) or mean (SD)
Mean age	50.2 (11.1)
Female	255 (91.7)
White	239 (86.0)
Concomitant rheumatic or musculoskeletal conditions	
Osteoarthritis	124 (44.6)
Fibromyalgia	85 (30.6)
Other rheumatic or musculoskeletal conditions	122 (43.9)
Mean years since RA diagnosis	9.4 (10.1)
Current RA treatment	
Biologics with/without csDMARDs	176 (63.3)
tsDMARDS with/without csDMARDS	34 (12.2)
csDMARDs without background tsDMARDs	68 (24.5)
Currently employed	154 (55.4)
Regular daytime shift work schedule (i.e., 9-5)	130 (46.8)
Mean daily Pain NRS	4.94 (2.46)
Mean daily Fatigue NRS	5.35 (2.52)

csDMARDS: conventional synthetic disease-modifying antirheumatic drugs; NRS: numeric rating scale; RA: rheumatoid arthritis; SD: standard deviation; tsDMARDS: targeted synthetic disease-modifying antirheumatic drugs

Primary Objective - PRO and Digital Sensor Measures

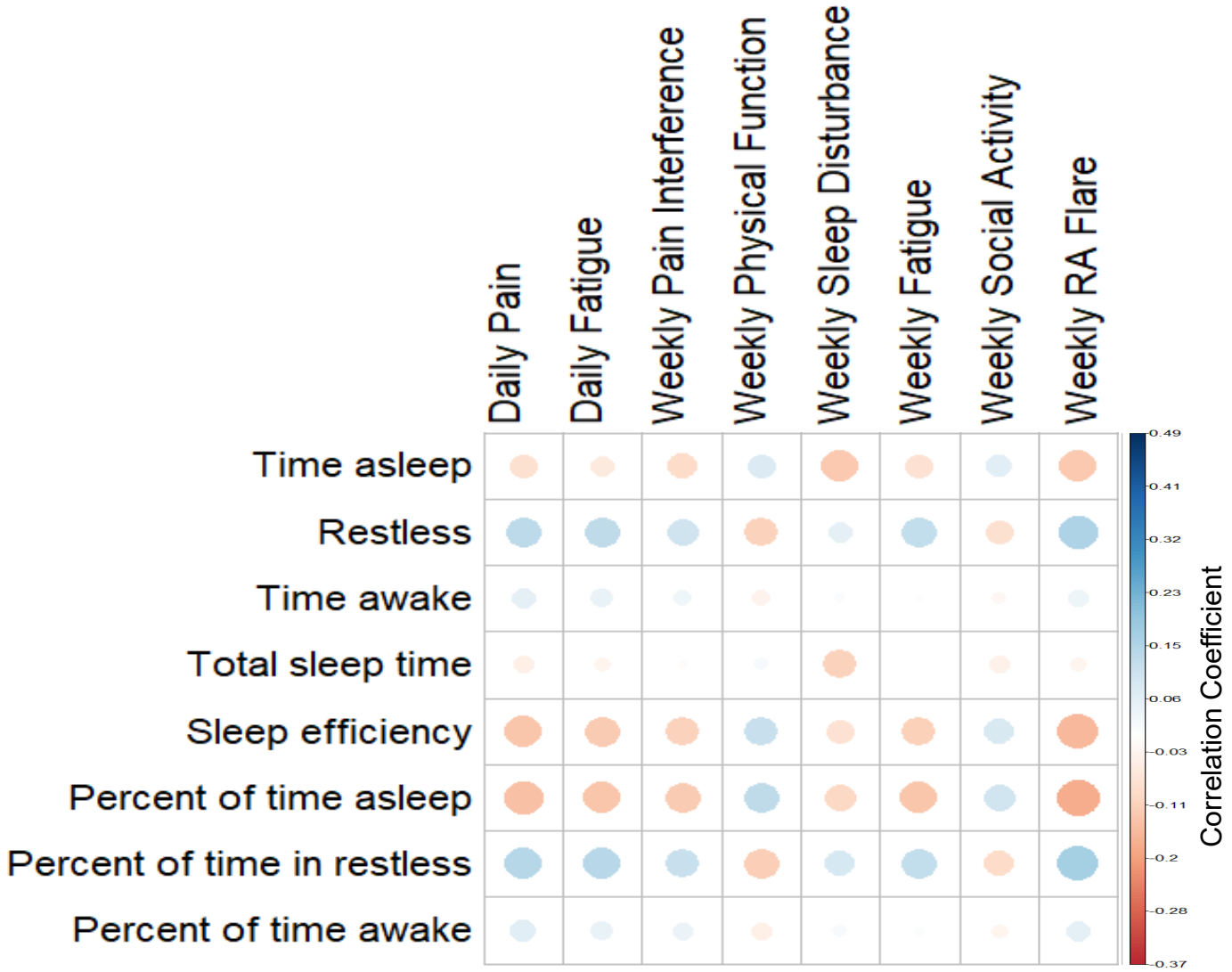


- Strongest correlation observed between step count/distance walked and patient-reported outcomes

- Derived activity measures not strongly correlated with patient-reported outcomes; derived data also includes manual activity tracking, not done by most participants
  - Algorithm for deriving activities is a black box

Note:
-Blue circles indicate direct relationship
-Red circles indicate inverse relationship
-Size of circles and depth of color reflect magnitude of association

Primary Objective - PRO and Digital Measures



- Strongest correlations observed between sleep efficiency and measures of pain/flare
- Difficult to draw conclusions of relationships between other derived sleep parameters and PRO measures

Note:
-Blue circles indicate direct relationship
-Red circles indicate inverse relationship
-Size of circles and depth of color reflect magnitude of association

Key Findings

- Low to moderate linear agreement with PRO measures and steps/distance – with greatest correlations observed with patient-reported physical function
- While not strongly correlated, associations between derived activity measures and patient-reported pain, fatigue, and physical function were significant
- In general, daily and weekly PRO measures were correlated in appropriate direction with smartwatch measures
- Very difficult to interpret sleep from Fitbit and sleep PRO measures
  - Sleep is less likely to be consciously adjusted by participants

Key Learnings

- Soft launch invaluable
- Participants were fairly adherent over 12 weeks
- Participants wanted daily reminders
- Algorithm black box makes interpretation of derived variables non-transparent/difficult
- Lead out period – participants dropped almost immediately without reminders, even with a consumer grade watch

Company Confidential ©2020 Eli Lilly and Company

Ongoing Analyses and Future Research Questions

Ongoing analyses:

- Reliability and validity of smartwatch measures
- Adherence and predictors of adherence with PRO measure completion and smartwatch use
- Classification of RA flares using digital data

Future research questions:

- Are intraday activity patterns relevant?
- How to interpret sleep from Fitbit and sleep PRO measures

Eliciting Meaningful Aspects of Physical Activity from Persons with Chronic Heart Failure

Milena Anatchkova, PhD
Senior Research Leader
Evidera, Inc.

Acknowledgments

- PRO Consortium Chronic Heart Failure Working Group

Company/Organization	Representative
Amgen	Gary Globe, PhD, MBA; Siddique A. Abbasi, MD, MSc
AstraZeneca	Folke Folkvaljon, MSc
Bayer	Luke Bamber, MSc
Bristol Myers Squibb	Brandon Becker, PhD; Shawn Li, PhD
Eli Lilly & Company	Jiat Ling Poon, PhD; Elizabeth Nicole Bush, MHS
GlaxoSmithKline	Linda Nelsen, MHS; Robyn von Maltzahn, MSc
J & J Janssen	Renee Pierson, MBA; Jeremiah Trudeau, PhD; John Whang, MD
Merck Sharp & Dohme Corp	Josephine Norquist, MS
Affiliation	**Other Participants**
Gwaltney Consulting	Chad Gwaltney, PhD
ePRO Consortium	Bill Byrom, PhD (Signant Health); Paul O'Donohoe, MSc (Medidata Solutions)
Universite' de Montre'al	Marc Jolicoeur, MD, MSc, MHS
Research Partner	**Research Team**
Evidera	Milena Anatchkova, PhD; Sonal Mansukhani, PhD; Saifra Sohail, MPH; Fanyang Zeng, BS; Ana Liberato, PhD

- The project described in this presentation is supported by the FDA of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award **U01FD006877-01** totaling $249,983 with 100% funded by FDA/HHS. The contents are those of the author and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS or the U.S. Government.

Background

- The US Food and Drug Administration's (FDA's) recent draft guidance on endpoints for drug development in heart failure (HF) states that an effect on symptoms or physical function, without a favorable effect on survival or risk of hospitalization, can be a basis for approving drugs to treat HF.[1]
- A gap exists in the literature on aspects of objectively quantifiable physical activity that are meaningful or important to persons with HF.
- Published recommendations on the selection and evaluation of wearable devices note the need for evidence that a device is measuring a meaningful aspect of the disease/condition or treatment from a patient's perspective.[2, 3, 4]
- The Patient-Reported Outcome (PRO) Consortium's Chronic Heart Failure (CHF) Working Group is pursing FDA qualification of a CHF activity monitor-based endpoint measure (DDT COA #000114).

[1] United States (US) Food and Drug Administration (FDA). Treatment for Heart Failure: Endpoints for Drug Development Guidance for Industry. 2019
[2] United States (US) Food and Drug Administration (FDA). Patient-Focused Drug Development: Collecting Comprehensive and Representative Input: Draft Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders. 2018;.
[3] United States (US) Food and Drug Administration (FDA). Patient Engagement in the Design and Conduct of Medical Device Clinical Investigations. 2019;
[4] Byrom B, Watson C, Doll H, et al. Selection of and Evidentiary Considerations for Wearable Devices and Their Measurements for Use in Regulatory Decision Making: Recommendations from the ePRO Consortium. *Value Health*. 2018;21(6):631-63

Objectives and Methods

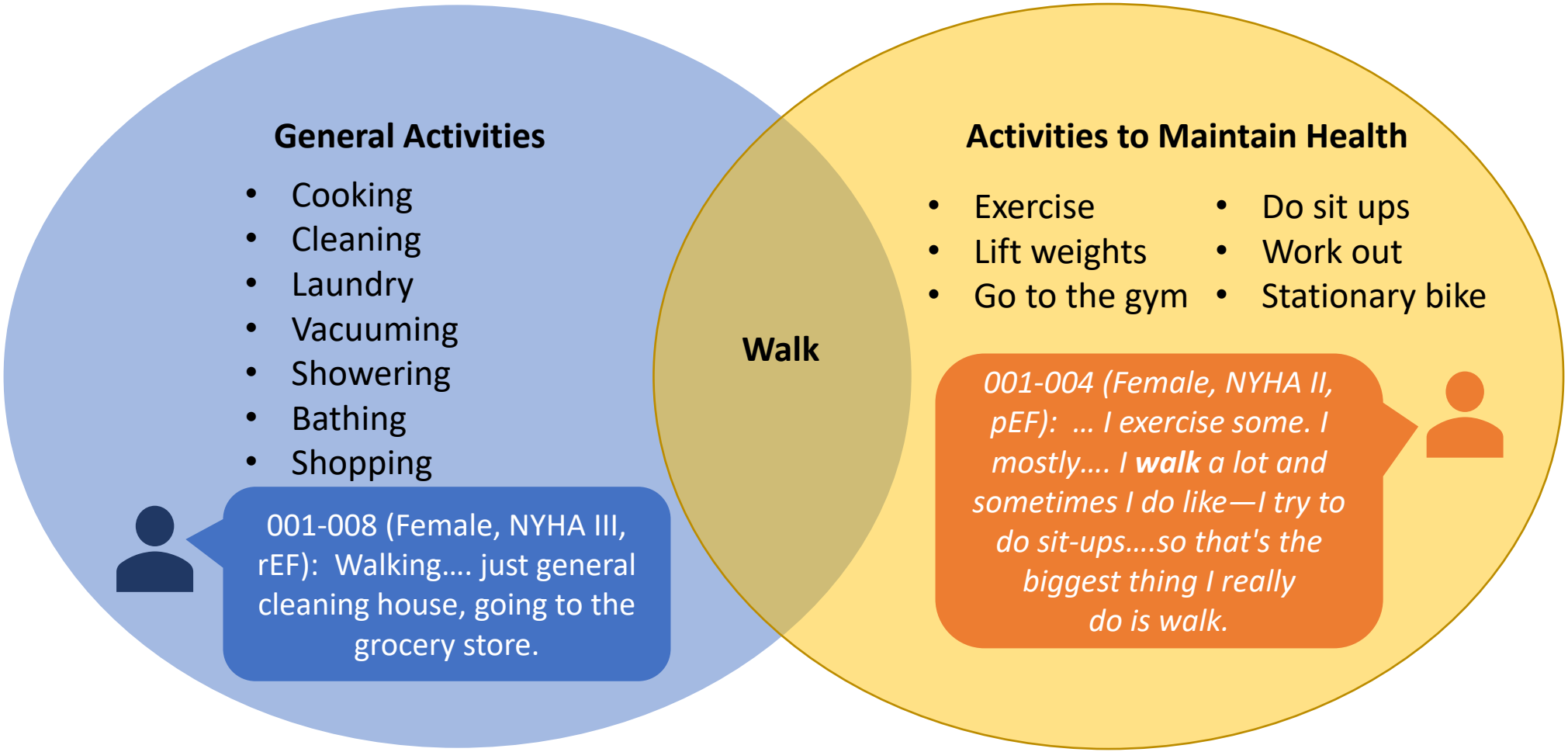
- **Objectives**
  - To characterize the types of physical activities, both inside and outside the home, that are important and meaningful to persons with CHF and then determine whether we can measure the essence of that physical activity with an activity monitor
- **Methods**
  - Stratified, non-interventional, qualitative interview study was conducted among persons with CHF.
  - Interviews were conducted via telephone and were audio-recorded.
  - Up to 40 participant interviews were planned.
  - Saturation was evaluated and demonstrated following the completion of the 28th interview.
  - Participant recruitment was stopped after the 31st interview.
- **Analysis**
  - Descriptive statistics
  - A coding dictionary was prepared based on the interview guide and reviewed by C-Path.
    - 24 codes for indoor activities; 14 had an eliciting probe in the interview
    - 22 codes for outdoor activities; 12 had an eliciting probe in the interview
  - A content-analysis approach was used to analyze qualitative data.

Participants

- 31 participants from 4 sites across the US (Tennessee, Michigan, Georgia, South Carolina)
- Median age: 68 years old
- 51.6% (n=16) were female
- 54.8% (n=17) were White
- 38.7% (n=12) reported using assistive device
- New York Heart Association (NYHA) HF severity:
  - Class II: n=15; 48.4%
  - Class III: n=12; 38.7%
  - Class IV: n=4; 12.9%
- CHF type:
  - HF with preserved ejection fraction (HFpEF): n=12 (38.7%)
  - HF with reduced ejection fraction (HFrEF): n=19 (61.3%)

Definition of Physical Activity

What does it mean to be physically active?



Physical Activity Inside the Home

- Saturation of concepts
  - 22/24 concepts (92%) of inside the home (or “indoor”) codes were utilized at least once by the 4th interview.
  - All codes were used by the 16th interview.
  - Unique new activities were not mentioned after the 21st interview, indicating saturation was reached.
  - New activities included: hobbies (n=3), dancing (n=2), playing with grandkids (n=2), feeding pets (n=1), and washing vehicle (n=1).
- Result patterns
  - Activities with elicitation probes were endorsed as relevant by most participants but did not consistently emerge spontaneously.
  - Spontaneously endorsed activities were primarily related to accomplishing household chores.
  - Household chores were also most commonly selected among the 3 most meaningful activities to participants along with walking.
  - Climbing stairs was the most commonly reported activity that participants avoided doing, followed by cooking and vacuuming.

Indoor Physical Activity: Concepts with interview guide probes

Indoor Physical Activities	Spontaneous Endorsements	Probed Endorsements	Frequency of being Endorsed as "Meaningful"	Does Activity Regularly	Can do but avoids doing	Doesn't do activity/Can't do
Taking a shower/ bath	6	25	15	6	1	1
Walking on flat surface	11	20	9	4	0	0
Getting dressed or undressed	4	27	6	3	0	0
Going up stairs	7	19	5	4	8	1
Toileting (being able to get on and off the toilet and perform personal hygiene)	1	30	4	1	0	0
Lifting items	1	27	3	1	0	2
Going down the stairs	6	22	2	2	5	1
Laying down flat on back	0	19	2	0	3	11
Bending	2	29	1	0	0	0
Getting out of bed	1	30	1	1	0	0
Standing for a few minutes at a time	3	27	1	0	0	1
Carrying items	2	29	0	4	0	0
Standing up from seated	4	27	0	1	0	0
Transferring (being able to get in and out of bed or a chair without assistance)	0	27	0	0	0	0

Indoor Physical Activity: Concepts with no probes in the interview guide

Indoor Physical Activities	Spontaneous Endorsements	Probed Endorsements	Frequency of being Endorsed as "Meaningful"	Does Activity Regularly	Can do but avoids doing	Doesn't do activity/Can't do
Cooking	19		16	12	5	3
General cleaning	22		16	8	3	6
Doing laundry	19		5	6	0	0
Vacuuming	16		4	5	4	1
Washing dishes	21		3	4	2	0
Picking up/ tidying up	15		2	1	0	1
Washing floors	12		1	0	3	3
Washing windows	3		1	0	1	3
Changing sheets	14		1	4	0	0
Throwing out the trash	5		1	0	0	0
Small repairs in the house ("reaching", climbing a ladder, painting jobs)	6		1	NA	NA	NA
Playing with grandkids*	2		1	NA	NA	NA
Feeding their pets*	1		1	NA	NA	NA
Dancing*	2		1	NA	NA	NA
Washing their vehicles*	1		0	NA	NA	NA
Making the bed	1		0	NA	NA	NA
Hobbies- (genealogy, building crafts, woodworking)*	3		0	NA	NA	NA

*Concept not included a priori in coding dictionary.

Dimensions of Indoor Activity

- For 3 indoor activities that participants noted as most meaningful, they were asked to rate the importance of the activity dimensions.
- The ability to do the activity and fulfill a task in general emerged as most relevant (mean rating across all indoor activities rated was 9), followed by intensity (mean rating across all indoor activities rated was 8.2).
- Speed was least relevant (mean rating across all indoor activities rated was 6.4).

Activity identified as one of 3 most important*	Dimension Importance** 0=Not important at all/ 10=Extremely important					
	Number of Participants	Speed	Duration	Frequency	Intensity	Fulfillment of Task
Cooking	16	6.3	7.1	6.8	7.5	9.1
Showering/bathing	14	6.8	6.7	8.6	8.7	9.3
General cleaning	13	5.2	7.9	7.7	8.3	9.2

*Only ratings for activities from >5 participants are presented.

**Some participants did not rate all dimensions.

Physical Activity Outside the Home

- Saturation of concepts
  - 22/22 (100%) of outside the home (or "outdoor") activity codes were utilized at least once by the 6th interview.
  - Unique new activities were not mentioned after the 22nd interview, indicating saturation was reached in the study.
  - New activities included: yard work/landscaping (n=12), grilling/barbecue (n=8), go to gym/sports (n=7), fishing (n=5), activities with animals (n=4), work on cars (n=2), and play with grandkids (n=2).
- Result patterns
  - A larger number of relevant outdoor activities not included a priori in the guide were mentioned by participants, but not all of them were noted as most meaningful.
  - Activities most frequently emerging spontaneously tended to also be selected as "most meaningful."
  - "Walking on flat surface" was the activity that was most often selected among the 3 most meaningful outdoor activities.
  - "Gardening" and "Walking on flat surface" were most often reported as activities participants don't do. The latter finding could be related to being unable to walk more than 10 minutes as probed.
  - Strenuous activities (e.g., hiking, jumping, going to the gym) were not selected as meaningful.

Outdoor Physical Activity: Concepts with interview guide probes

Outdoor Physical Activities	Spontaneous Endorsements	Probed Endorsements	Frequency of being Endorsed as "Meaningful"	Does Activity Regularly	Can do but avoids doing	Doesn't do activity/ Can't do
Walking on flat surface *	15	16	17	8	3	7
Shopping	16	13	11	8	3	3
Gardening	16	7	9	2	4	9
Running errands	10	15	7	4	3	5
Doctor visits	5	23	7	3	0	0
Bank	3	24	3	3	0	4
Walking on inclined surface	2	24	2	1	4	4
Going up stairs	2	23	0	0	4	2
Going downstairs	2	24	0	0	4	0
Carrying items	0	25	0	0	1	3
Lifting items	0	22	0	0	1	3
Bending	0	21	0	0	3	1
Standing	0	22	0	1	0	2

Activities most frequently emerging spontaneously tended to also be selected as "meaningful."

Other anticipated activities were noted as relevant but were endorsed less often as most meaningful.

*The probe was for walking >10 minutes

Outdoor Physical Activity: Concepts with no interview guide probes

Outdoor Physical Activities	Spontaneous Endorsements	Probed Endorsements	Frequency of being Endorsed as "Meaningful"	Does Activity Regularly	Can do but avoids doing	Doesn't do activity/Can't do
Attending events	6		6	2	1	3
Yard work/landscaping*	12		6	NA	NA	NA
Driving	15		5	0	3	2
Friends/Families visits	8		4	1	0	0
Swimming	2		2	0	2	2
Fishing*	5		2	NA	NA	NA
Activities with Animals (Birds and dogs)*	4		2	NA	NA	NA
Biking	3		1	0	2	5
Jogging	1		1	0	0	3
Dancing	1		1	1	0	0

*Concept not included a priori in coding dictionary.
Note: Additional concepts that emerged as relevant but not selected as most meaningful include: Going to church* Checking mail*, Go to gym/sports*, Work on cars*, Hiking, Jumping, Play with grandkid*, Grilling/ barbecuing* Hunting*, Traveling*

Dimensions of Outdoor Activity

- Pattern of results for ratings of dimensions of outdoor activities was similar.
- The ability to do the activity and fulfill a task in general emerged as most relevant (mean rating across all outdoor activities rated 8.7).
- Speed was least relevant (mean rating across all outdoor activities rated 6.3).

Activity identified as one of 3 most important*	Dimension Importance** 0=Not important at all/ 10=Extremely important					
	Number of Participants	Speed	Duration	Frequency	Intensity	Fulfillment of Task
Walking on flat surface	17	5.4	8	8.2	8.1	9.2
Going shopping	9	6.8	6.2	6.9	7.9	8.8
Going to the doctor/dentist	8	8.4	7.6	7.4	5.9	8.1
Gardening	6	7.6	7.5	9.3	9.2	9.3

*Only ratings for activities from >5 participants are presented.

**Some participants did not rate all dimensions.

Walking Indoors is Different Than Walking Outdoors

Activities explored and noted as relevant both indoors and outdoors

Walking on flat surface more important (17 vs. 9) and more difficult (7 vs. 0) outdoors

Indoor Physical Activities	Spontaneous Endorsements	Probed Endorsements	Endorsed as "Most Meaningful"	Does Activity Regularly	Can do but avoids doing	Doesn't do activity/Can't do
Walking on flat surface	11	20	9	4	0	0
Going up stairs	7	19	5	4	8	1
Going down the stairs	6	22	2	2	5	1
Lifting items	1	27	3	1	0	2
Carrying items	2	29	0	4	0	0
Outdoor Physical Activities						
Walking on flat surface	15	16	17	8	3	7
Going up stairs	2	23	1	0	4	2
Going downstairs	2	24	1	1	4	0
Lifting items	0	22	0	0	1	3
Carrying items	0	25	0	0	1	3

Stairs, lifting, and carrying items more relevant indoors

Lessons Learned

- Content-related lessons:
  - Participant narratives focus on physical activities related to accomplishing tasks or relating to others, less so on individual activity components captured by activity monitor devices.
  - Walking is a common construct readily measured by devices and recognized as important by participants both indoors and outdoors.
  - Some translation of activity monitor constructs to activities noted as meaningful and dimensions noted as important to people with CHF may be needed.
  - Context impacts perceptions of relevance and importance of physical activities.
- Design-related lessons:
  - From the participant perspective, the concepts of physical function, physical activity, activities of daily living, and participation have substantial overlap.
  - While activity monitors allow for a very detailed evaluation of different dimensions of activity, there were notable challenges in discussing these activity dimensions with participants, especially for task-oriented activities (e.g., going to the doctor).
- Pandemic-related challenges:
  - Logistics
  - Changes in level of activity associated with pandemic restrictions may have impacted responses.

Panel Discussion

Moderator

- *Maria Mattera, MPH* – Assistant Director, PRO Consortium, C-Path

Presenters

- *Jennifer Goldsack, MChem, MA, MBA* – Executive Director, Digital Medicine Society (DiMe)
- *Bill Byrom, PhD* – Vice President, Product Intelligence and Positioning, Signant Health
- *Jiat Ling Poon, PhD* – Principal Research Scientist, Eli Lilly and Company
- *Milena Anatchkova, PhD* – Senior Research Leader, Evidera, Inc.

Additional Panelists

- *Andrew Potter, PhD* – Mathematical Statistician, Division of Biometrics I, Center for Drug Evaluation and Research, U.S. Food and Drug Administration
- *Steven Blum, MBA, MA* – Asset and Indication Lead, Patient-Reported Outcomes Assessment, WWHEOR, Bristol Myers Squibb