Endpoint Construction from Activity Monitor Data: Chronic Heart Failure

Tenth Annual
Patient-Reported Outcome Consortium Workshop

April 24 – 25, 2019 ■ Silver Spring, MD

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 Outline

- Measuring physical activity with wearable devices in Chronic Heart Failure (CHF) trials: Addressing unmet needs
- Overview of Chronic Heart Failure Working Group
- An industry perspective
- An FDA perspective
- Proposed activity-based endpoints for CHF using activity monitors
- Panel Discussion
- Question and Answer

Session Participants

Moderator

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

Presenters

- *Chad Gwaltney, PhD* – President, Gwaltney Consulting
- *Maria Mattera, MPH* – Assistant Director, PRO Consortium, C-Path
- *Jeremiah (Jay) Trudeau, PhD* - Director, Patient-Reported Outcomes, Janssen Global Services
- *Wayne Amchin, RAC, MIA, MPA* – Senior Consumer Safety Officer, Division of Cardiovascular and Renal Products, OND, CDER, FDA
- *Bill Byrom, PhD* – Vice President of Product Strategy and Innovation, CRF Bracket

Panelist

- *Ebony Dashiell-Aje, PhD* – Reviewer, COA Staff, OND, CDER, FDA

Measuring Physical Activity with Wearable Devices in CHF Trials: Addressing Unmet Needs

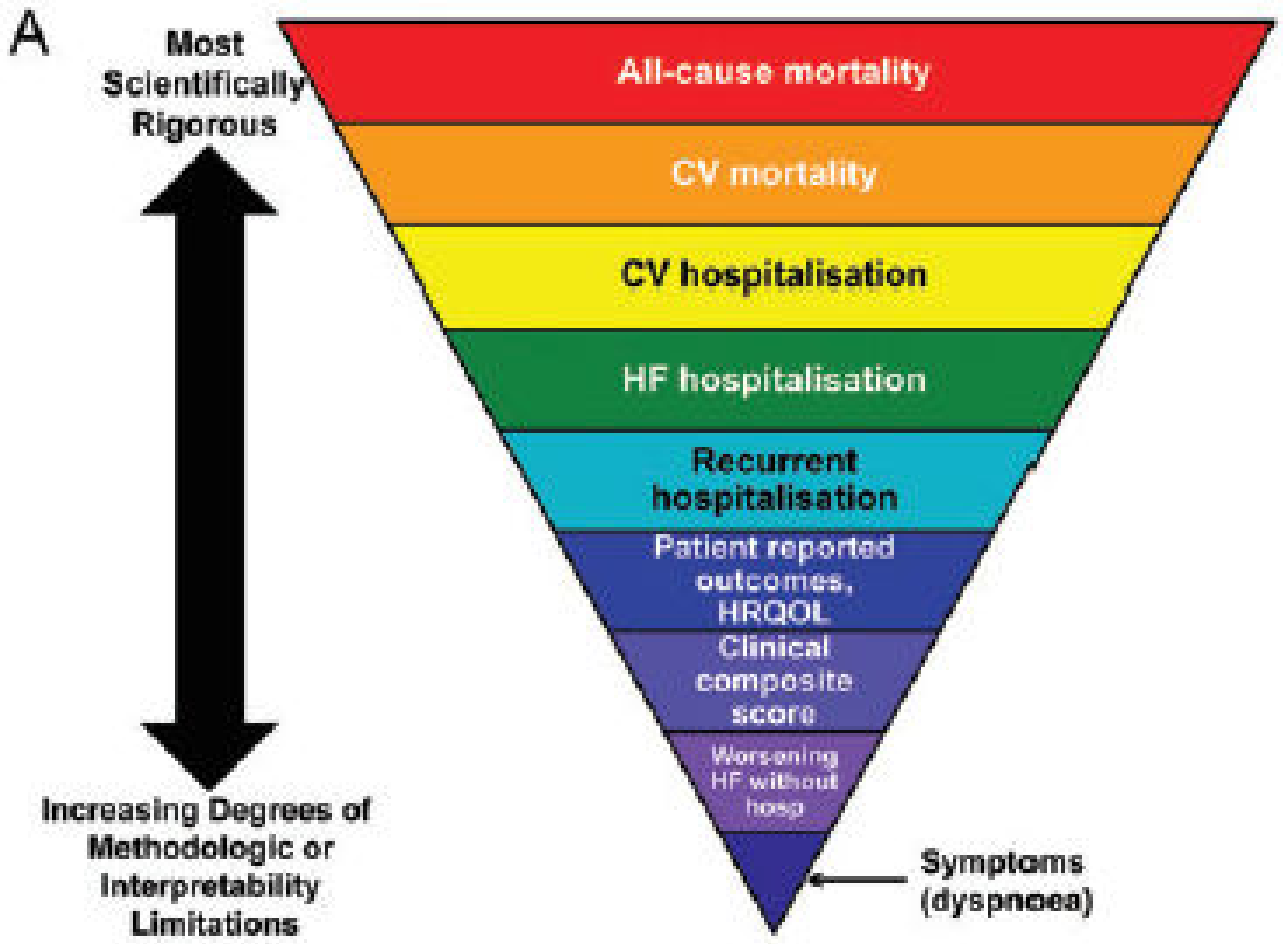
Chad Gwaltney, PhD – President, Gwaltney Consulting

Heart Failure: Significant Global Burden

- Chronic Heart Failure (CHF) = A shared chronic phase of many cardiac diseases[1,2]
  - "A complex clinical syndrome that results from any structural or functional impairment of ventricular filling or ejection of blood"
- Estimated that 37.7 million people globally are living with CHF; 1-2% of adult population in developed countries[1]
- Prevalence of CHF increases with age: An "epidemic" as the global population ages[1]
  - By 2030, US prevalence expected to increase by 46%[3]
- Despite improvements in treatment, heart failure is still a lethal condition
  - 1- and 5-year mortality rates of approximately 20 and 53%, respectively[4]
  - May vary across subgroups (HFrEF, HFpEF)[5]

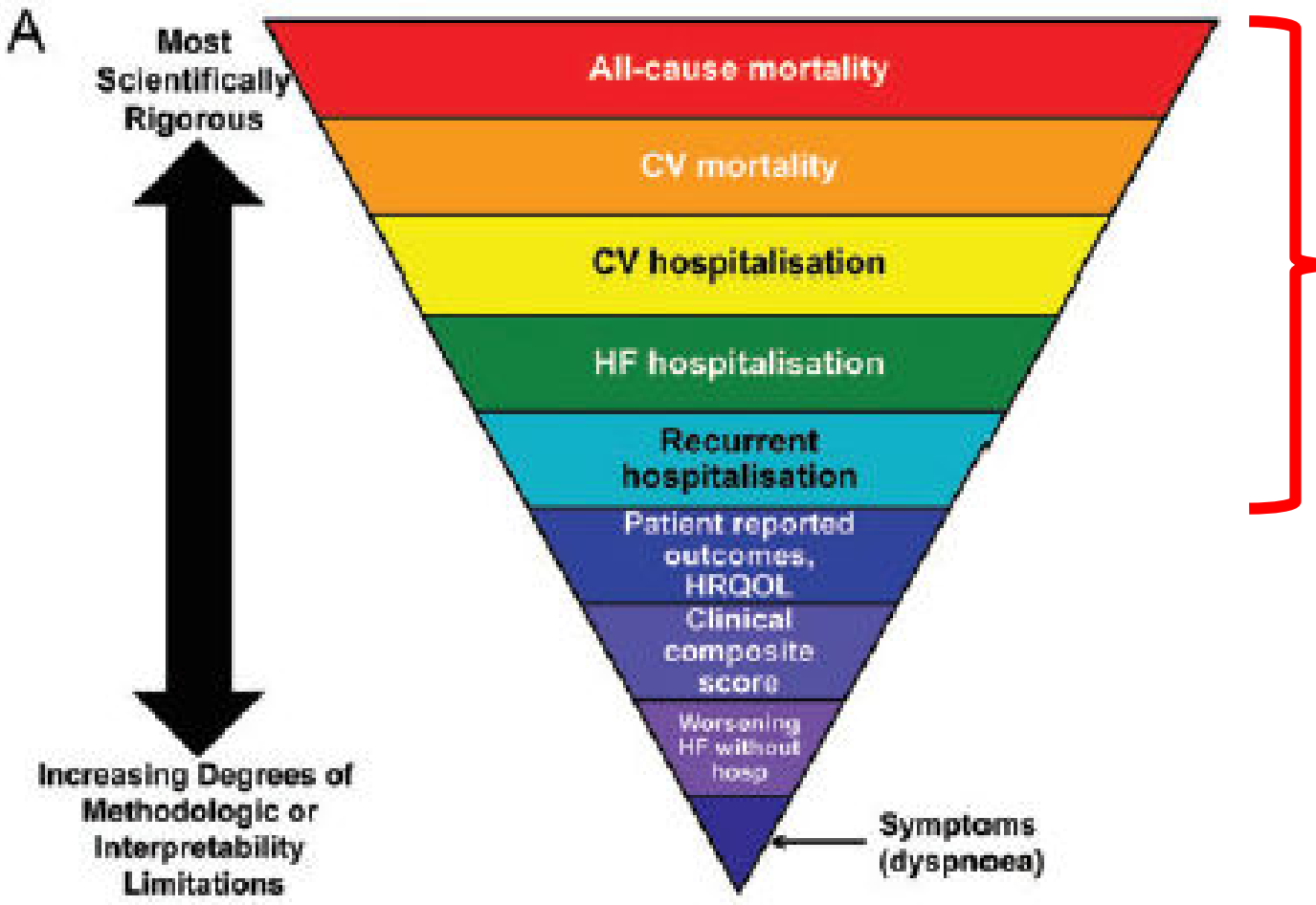
[1] Ziaeian & Fonarow 2016; [2] Yancy 2013; [3] Heidenreich 2013; [4] Gerber 2015; [5]Lam 2018

Endpoints in CHF Trials



Zannad 2015

Endpoints in CHF Trials



- Requires lengthy trials
- Requires large sample size
- May not be right outcome for all subgroups
- Does not capture patient's everyday experience

Need for Novel Endpoints

Development of Therapeutics for Heart Failure

Exploring New Endpoints for Patients With Heart Failure With Preserved Ejection Fraction

"Wearable devices are evolving rapidly and may be helpful for monitoring patient activity."

Butler et al. Exploring New Endpoints for Patients With Heart Failure With Preserved Ejection Fraction. Circ Heart Fail. 2016; 9: e003358

Symptoms and Physical Activity Are Highlighted in ClinRO CHF Measures

NYHA class	Symptoms
I	No limitation in normal physical activity
II	Mild symptoms only in normal activity
III	Marked symptoms during daily activities, asymptomatic only at rest
IV	Severe limitations, symptoms even at rest

NYHA: New York Heart Association.

Patient- and provider-determined NYHA class are poorly correlated (r=0.40) and patient ratings are consistently lower than clinician ratings (Williams 2017); poor inter-rater agreement (Raphael 2007)

Use of Activity Monitors To Capture Physical Activity in CHF Patients

- Many sponsors interested in implementing activity monitors in clinical trials, but there are still many unanswered questions
  - Which activity monitor(s) to use?
  - Which physical activity variables are the most relevant for CHF patients?
  - How to establish the reliability and validity of scores from activity monitors?
  - How to aggregate data?
  - How to handle missing data? What does missing data look like?
  - How to create an endpoint that can be used in a trial?
  - What is a threshold for meaningful change in physical activity?

CHF Working Group Addresses Significant Unmet Need

- Excellent opportunity for a collaborative effort among sponsors who are attempting to answer similar questions regarding the use of activity monitors
- Value specific to CHF: WG will address issues related to the use of activity monitors to measure physical activity and creation of endpoints to evaluate treatment in CHF trials
- Value beyond CHF: Same issues are relevant in other therapeutic areas where physical activity is a relevant outcome

References

- Butler et al. Exploring New Endpoints for Patients With Heart Failure With Preserved Ejection Fraction. *Circ Heart Fail.* 2016; 9: e003358.
- Gerber Y, Weston SA, Redfield MM, et al. A contemporary appraisal of the heart failure epidemic in Olmsted County, Minnesota, 2000 to 2010. *JAMA Intern Med*. 2015;175(6):996-1004.
- Heidenreich PA, et al. Forecasting the impact of heart failure in the United States a policy statement from the American Heart Association. *Circ. Heart Fail.* 2013; 6:606–619.
- Lam CSP et al. Heart failure with preserved ejection fraction: from mechanisms to therapies. *Eur Heart J.* 2018; 39: 2780–2792.
- Raphael C, Briscoe C, Davies J, et al. Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. *Heart* 2007;93:476-482.
- Williams BA et al. Agreement between heart failure patients and providers in assessing New York Heart Association functional class. *Heart & Lung: The Journal of Acute and Critical Care*. 2017; 46: 293 – 299.
- Yancy CW et al. 2013 ACCF/AHA Guideline for the Management of Heart Failure. *Circulation.* 2013;128:e240-e327.
- Zannad F et al. Clinical outcome endpoints in heart failure trials: a European Society of Cardiology Heart Failure Association consensus document. *European Journal of Heart Failure*. 2013; 15: 1082–1094.
- Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol*. 2016;13(6):368-78.

Chronic Heart Failure Working Group: Overview

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

Background

Rationale for Chronic Heart Failure (CHF) Working Group

- CHF was identified as a priority area for development of a more patient-centered approach to the evaluation of clinical benefit in CHF treatment trials.
- Based on the increased availability of mobile sensor technologies (e.g., activity trackers/monitors), there is substantial interest in leveraging these to assess novel clinical trial endpoints.
- During formation of the Working Group, Amgen offered to share its developmental PRO measures and results of ongoing efforts supporting use of activity monitor data in patients with CHF.

Goals of the CHF Working Group

- Develop a measurement strategy for the assessment of symptom severity, symptom impact on physical function, and physical activity for adults with CHF by incorporating both patient-reported and activity monitor data
- Obtain FDA qualification of this combination of measures to assess efficacy endpoints in CHF clinical trials

Concepts of Interest

- The concepts of interest for the PRO measures, which were developed by Amgen, are:
  - self-reported severity of chronic heart failure symptoms (*Chronic Heart Failure-Symptom Scale* [*CHF-SS*])
  - self-reported impact of chronic heart failure symptoms on physical functioning (*Chronic Heart Failure-Impact Scale* [*CHF-IS*])
- The concept of interest for the activity monitor-based endpoint measure is physical activity with specific variables to be determined.

Context of Use

The target population includes adults…

- with clinician-confirmed CHF for ≥3 months with New York Heart Association class II–IV symptoms for ≥4 weeks,
- with a diagnosis of CHF with preserved ejection fraction (HFpEF) or CHF with reduced ejection fraction (HFrEF),
- in stable condition for at least 4 weeks, and
- treated with stable, optimal pharmacological therapy for a minimum of 4 weeks prior to screening.

The scores from the proposed PRO measures and activity monitor-based endpoint measure will be used to derive efficacy endpoints in CHF treatment trials.

Brief Summary of Measures

- *Chronic Heart Failure-Symptom Scale* (*CHF-SS*)
  - 9 items addressing self-reported severity of CHF symptoms
  - 7-day recall period
  - 5- or 6-level verbal rating scale
- *Chronic Heart Failure-Impact Scale* (*CHF-IS*)
  - 9 items addressing self-reported impacts of CHF symptoms on physical functioning
  - 7-day recall period
  - 6-level verbal rating scale
- Activity Monitor-Based Endpoint Measure
  - Physical activity monitoring device variable(s) (to be determined)
  - Agnostic to device

Example Endpoint Model for Treatment of CHF

Endpoint Hierarchy	Endpoint Concept(s)	Endpoint Type
Primary	Time to cardiovascular (CV) death or time to heart failure (HF) event	Event rate
Secondary	Evaluate effects of [*Drug X*] on time to: • CV death • HF hospitalization • All-cause death	Event rate
Potential New Primary or Secondary	• Reduction in (or delayed worsening of) severity of CHF symptoms • Reduction in (or delayed worsening of) limitations in physical function • Improvement in (or delayed worsening of) activity monitor-based variable reflecting a meaningful aspect of physical activity/mobility	• PRO • PRO • Activity monitor-based COA

Completed Activities

- Prior research completed by Amgen on the *CHF-SS* and the *CHF-IS* in both HFpEF and HFrEF patients confirmed item relevance, concept coverage, and appropriateness of response options and recall period.

- FDA feedback was provided to Amgen on several occasions during the development process.
  - In the latest communication, FDA requested additional qualitative evidence from HFpEF and HFrEF patients.

- Amgen has agreed to share these measures with the CHF Working Group for qualification.

- Letter of Intent was submitted to FDA in December 2018.

Unique Issues for the Working Group

- This is the PRO Consortium's first working group that is proposing qualification of an activity monitor-based endpoint measure.

- One of the main challenges is determining what variable(s) from the activity monitor will reflect a sufficiently meaningful aspect of physical activity and, therefore, used to derive the endpoint.

- It remains an empirical question as to whether it makes clinical and psychometric/ clinimetric sense to combine the PRO data with activity monitor-based data to derive a composite endpoint.

Next Steps

- Additional cognitive interviews with the *CHF-SS* and *CHF-IS* are being conducted by Amgen to obtain the additional qualitative evidence requested by FDA; patient recruitment is already under way.
- Further psychometric evaluation of the measures is planned as part of Amgen's CHF development program.
- A stand-alone study is also planned by Amgen to evaluate the use and usefulness of an activity monitor in CHF treatment trials, including evaluation of the data to identify the variables that would support an endpoint.
- Additional qualitative research is needed to provide background on what variable(s) would reflect a meaningful aspect of physical activity to patients that could be derived from the activity monitor.
  - Several sponsors have indicated their willingness to share qualitative evidence in CHF associated with this topic.

Working Group Sponsors

- Amgen
- AstraZeneca
- Bayer
- Ironwood Pharmaceuticals, Inc
- Janssen Global Services LLC
- Eli Lilly and Company
- Merck Sharpe & Dohme Corp.
- Novartis Pharmaceutical Corporation
- Sanofi

Conceptual Distinctions for Passive Monitoring and Patient-Reported Outcome Assessment

Jeremiah (Jay) Trudeau, PhD - Director, Patient-Reported Outcomes, Janssen Global Services

No substitute for patient perspectives

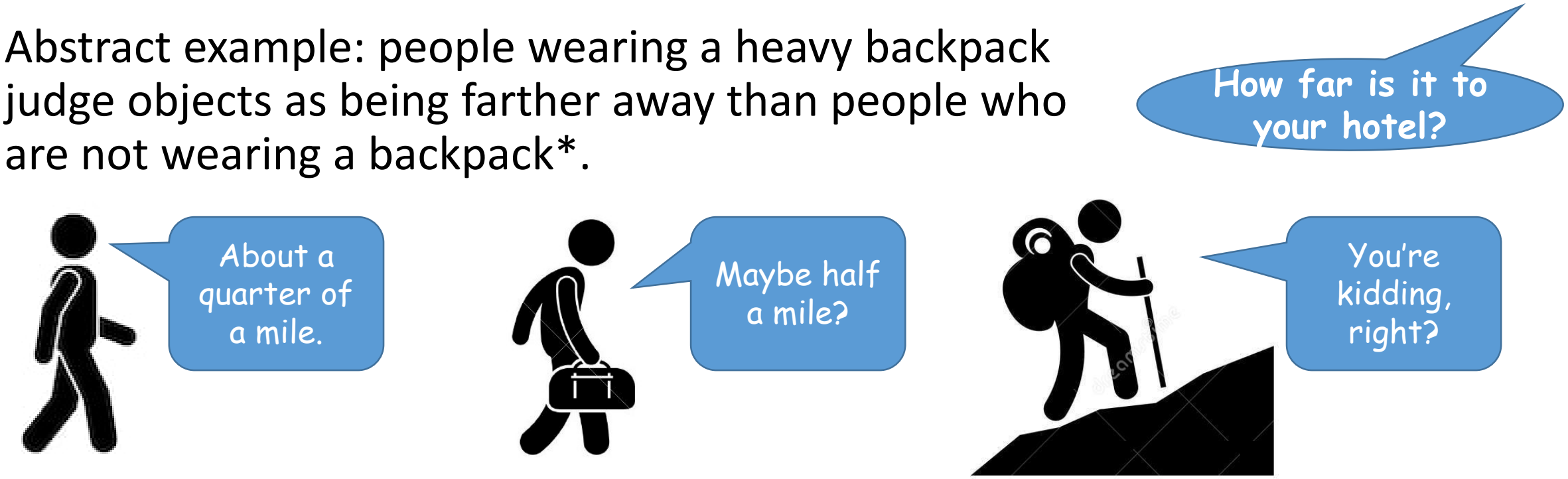
- Electronic platforms for ePRO and passive monitoring of patients generate fundamentally different types of data
  - Symptoms and feelings are inherently subjective; they cannot be observed or reported by anyone other than the patient
  - Observable behavior (signs) may be recorded objectively; they report things that happen, but not how they are perceived or experienced
- Subjective is no more or less valid than objective, just different
  - Drugs must demonstrate benefit to how patients ‘feel and function’
  - FEEL is inherently subjective; FUNCTION may be either
  - Patient reports may have more factors but better reflect what is intended
  - Objective reports have narrower scope but may measure different concepts

Only patients can tell us how patients feel.

Kubrick S (1968) 2001: A Space Odyssey; Metro-Goldwyn-Mayer.

But patient perceptions can be flawed

Abstract example: people wearing a heavy backpack judge objects as being farther away than people who are not wearing a backpack*.



Sometimes we may be better off using a different tool if what we want to measure is objectively observable.

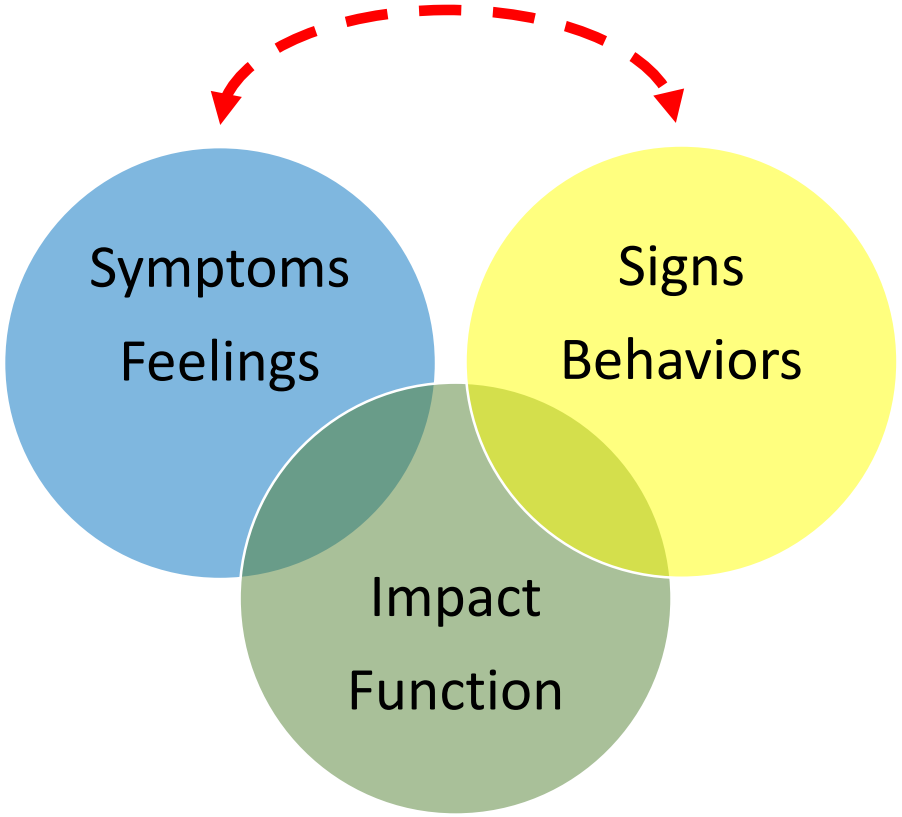
Less abstract example: CHF patients with high fatigue have lower maximal oxygen uptake (peak VO_2)**

*Proffitt DR, Stefanucci J, Banton T, Epstein W. The role of effort in perceiving distance. Psychological Science. 2003;14:106–112. **Evangelista LS, Moser DK, Westlake C, Pike N, Ter-Galstanyan A, Dracup K. Correlates of fatigue in patients with heart failure. Prog Cardiovasc Nurs. 2008;23(1):12-17.

Distinct concepts of interest

Observable ← - - → **Introspective**

(suitable for monitors)	*(suitable for PROs)*
• Sleep quantity	• Sleep quality
• Physical activity	• Physical function
• Heart rate	• Fatigue
• Achievement	• Effort



Related but conceptually distinct

Clarity of concept is vital for successful measurement!

What do wearables *really* measure?

- Passive data collection measures physical events, but does not proximally tell us how a patient *feels* or (usually) *functions*
- For example, a passive monitor might tell us:
  - Exactly when a patient's heart rate increases, but not if they are excited or in pain
  - How much a patient is moving around, but not if they are functioning or productive
  - Time spent asleep, but not how well they sleep or how rested they feel
- To assess patients' **experience** we must ask them
- To assess actual **behavior** we can monitor them
- We may want to infer feeling from observation, but we aren't measuring it

Wearable data are a complement to PRO data, not a substitute.

Summary

- Patient reports are the only way to directly assess patients' experience, but not the only way to measure what they do
- Subjective and objective data complement/supplement but cannot replace each other
- Innovative tools do not change the underlying principles of COAs
- New types of endpoints based on passive monitoring require clarity on concepts of interest. What is really patient-relevant?

Koan of the Day: If you take a step without your FitBit, did you really move?

U.S. FOOD & DRUG ADMINISTRATION

Do the endpoint(s) justify the means (Digital/Mobile technology):

An FDA perspective

U.S. FOOD & DRUG ADMINISTRATION

Disclaimer

This speech reflects the views of the author and should not be construed to represent FDA's views or policies.

Overview of my talk

- Endpoints, generally
- Endpoints for Heart Failure and Applicability of Wearables
- Approvability Considerations
- When to engage with the FDA about incorporating this technology in your development program

Endpoints

- Clinical investigation endpoints used to support labeling claims must be based on a well-defined and reliable assessment (see 21 CFR 314.126).

- An endpoint is a precisely defined variable intended to reflect an outcome of interest that is statistically analyzed to address a particular research question.

- A precise definition of an endpoint typically specifies the type of assessments made, the timing of those assessments, the assessment tools used, and possibly other details, as applicable, such as how multiple assessments within an individual are to be combined.

- Clinically relevant endpoints typically reflect how patients feel, function, or survive.

Endpoints

- When a mobile technology tool is used to collect data the clinical relevance of the data and endpoint should be assessed.

- When the tool is used as an alternative way to measure a validated disease outcome (e.g. Blood pressure), the traditional endpoint for that disease may be used.

- Sponsors intending to use mobile technology in a clinical investigation intended for submission as part of a marketing application should always discuss those proposed protocols with the relevant review division for review and comment prior to commencing such a trial.

Treatments for Heart Failure: Endpoints for Drug Development

- The agency is working on guidance on endpoints for drug development for the treatment of heart failure, including areas where the agency would like further discussion, for example symptomatic benefit. This planned guidance is not focused on wearables, but some principles are germane.

- The type of evidence of effectiveness needed to support approval of drugs for heart failure does not differ from the evidence needed to support the approval of drugs to treat other conditions: substantial evidence demonstrating that the drug improves how a patient feels or functions, i.e., symptomatic or functional improvement, or survives.

Treatments for Heart Failure: Endpoints for Drug Development

- Treatments that effect physical function can be a basis for approval of drugs to treat heart failure, without necessarily also demonstrating a favorable effect on survival or risk of hospitalization.

- Wearable technology can provide data on the effect on physical function and some symptoms, e.g., dyspnea.
  - drugs that improve symptoms or function when added to the SOC would be valuable, without necessarily improving survival or hospitalization
  - Drugs that provide substantial and persistent improvements in symptoms or function, for certain patients with some decrease in survival

Treatments for Heart Failure: Endpoints for Drug Development

- With wearables, still need:
  - Pre-specified endpoints
  - Statistical significance
  - To decide the treatment effect necessary for approval or to add a claim
- Known for endpoints we have used, e.g., 6MWD
- Unknown for wearable-measured endpoints where the endpoint has not been the basis for approval, e.g., 6MWD is known, required improvement in steps is not known

Efficacy Endpoints Related to How Patients Feel and Function

- For drugs intended for chronic use, wearables can help to assess durability of effect.
- Endpoints acceptable to FDA, that may be measurable with wearables-exercise capacity (e.g., 6MWD), functional capacity, and measures of physical function in activities of daily living.
- Wearable activity endpoints that may be acceptable to FDA but for which the required treatment effect has not been established or has not been the basis of approval-exercise achievement (total steps, total walk distance) at baseline vs. after treatment.

Exercise duration endpoint: modification opportunities with mobile/digital

- Mobile exercise measurement:
- Opportunity to measure functional capacity in the context of ADLs (physical function).
  - Are there times of day that a six-minute walk is a goal and other times of day it is less meaningful or not relevant? Mobile technology may provide opportunities to assess that.
  - If a six-minute walk is not the daily goal of patients, what exercise measures may be better clinical endpoints?
  - Is there an opportunity to measure exercise capacity in different kinds of weather?

Does the agency require 6MWD or Cardio-Pulmonary Exercise Testing (CPET)?

- If improvement in six-minute walk distance (6MWD) was marginal, but steps improved significantly, would that result in approval?

- If steps declined, but 6MWD improved in line with previous treatment effects, would that lead to approval?

Does the agency require 6MWD or CPET?

- If 6MWD was not used in the trial, but steps or some other measure of activities of daily living was used, would the agency consider such an application?
  - Pre-specified endpoint, proposed treatment effect, statistically significant result?
  - Yes, need to discuss with agency in advance.

Is the choice all traditional or all mobile/digital?

- Discuss with the relevant review division in OND ways you can incorporate this technology into your development program
- The earlier you incorporate the technology in your program, the better positioned you will be to use it in a pivotal trial
- To validate a particular technology, consider measurement of healthy volunteers using the traditional approach AND measure the same using the technology of interest
- Consider the potential of a study arm using this technology tied to a potential claim. Discuss this with the review division

When Should Sponsors Approach the Agency About Incorporating Digital/Mobile Technology in Development Programs?

- During your initial pre-IND meeting to have a conversation about how you potentially could incorporate these tools

- At EOP1 meetings, to discuss your phase 2 (P2) trial to leverage this technology in your dose-ranging studies

- At EOP2 meetings, to get agency input on use of this technology in your phase 3 (P3) pivotal trial, and, if you were forward thinking, how your use of this technology in P2 trials will have an impact on the design of your P3 trials.

- Consider taking advantage of the FDA's Type C Clinical Outcome Assessment Meeting and other related opportunities

Contact Information

- Wayne Amchin, RAC

  Senior Consumer Safety Officer

  Division of Cardiovascular and Renal Products

  Wayne.Amchin@fda.hhs.gov

  301-796-0421

  Focus:

  – Pulmonary Arterial Hypertension
  – Use of Digital and Mobile Technology in Clinical Trials
  – Development of Mobile App for PAH and Subsequently Repurposing the App for Other Therapeutic Areas with Related Endpoints and App Measurable Elements


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Activity-based Endpoints for CHF Using Activity Monitors

Bill Byrom, PhD – Vice President of Product Strategy and Innovation, CRF Bracket

@billbyrom

Measuring meaningful aspects of health

- **Symptoms** of CHF (shortness of breath, chest pain, fatigue) significantly impact ability to perform **activities of daily living**, and impact quality of life

  (Heo S. et al. Heart and Lung 2009; 38:100-108).

- **Average steps** per day in CHF patients: 4,342 steps/day

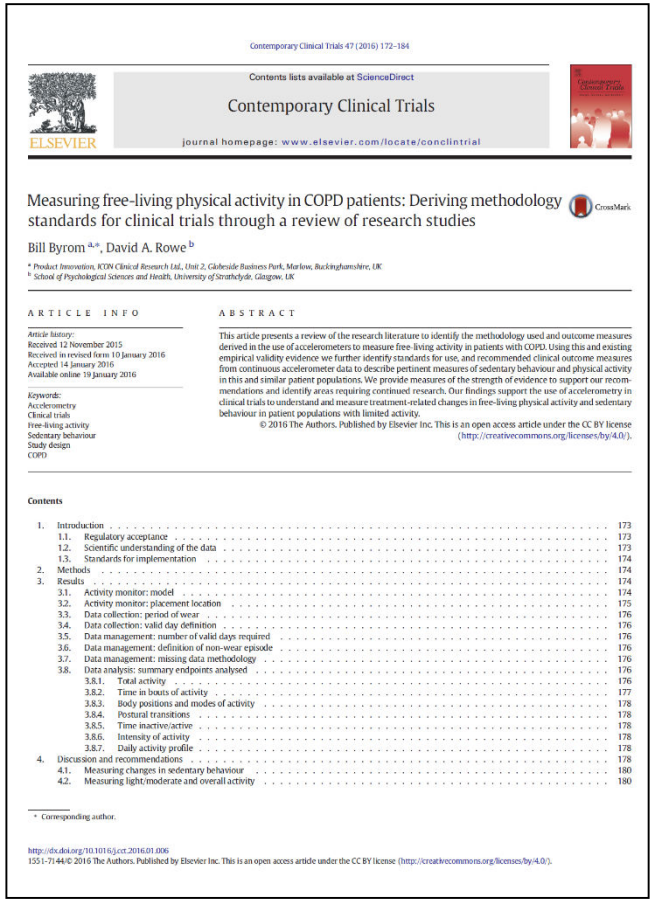
  (Houghton AH et a;. Eur. J. Heart Fail. 2002; 4: 289-295)

- **Six-minute walking test** distances:
  - NYHA II: 353 ± 91 m
  - NYHA III: 172 ± 89 m

  (Peeters and Mets. J Gerontology: Med Sci 1996; 51: M147-M151)

Measuring meaningful aspects of health

Byrom B and Rowe DA. Measuring free-living physical activity in COPD patients: Deriving methodology standards for clinical trials through a review of research studies. *Contemporary Clinical Trials*, 2016; **47**:172-184.



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.

Contents

* Corresponding author.

http://dx.doi.org/10.1016/j.cct.2016.01.006
1551-7144/

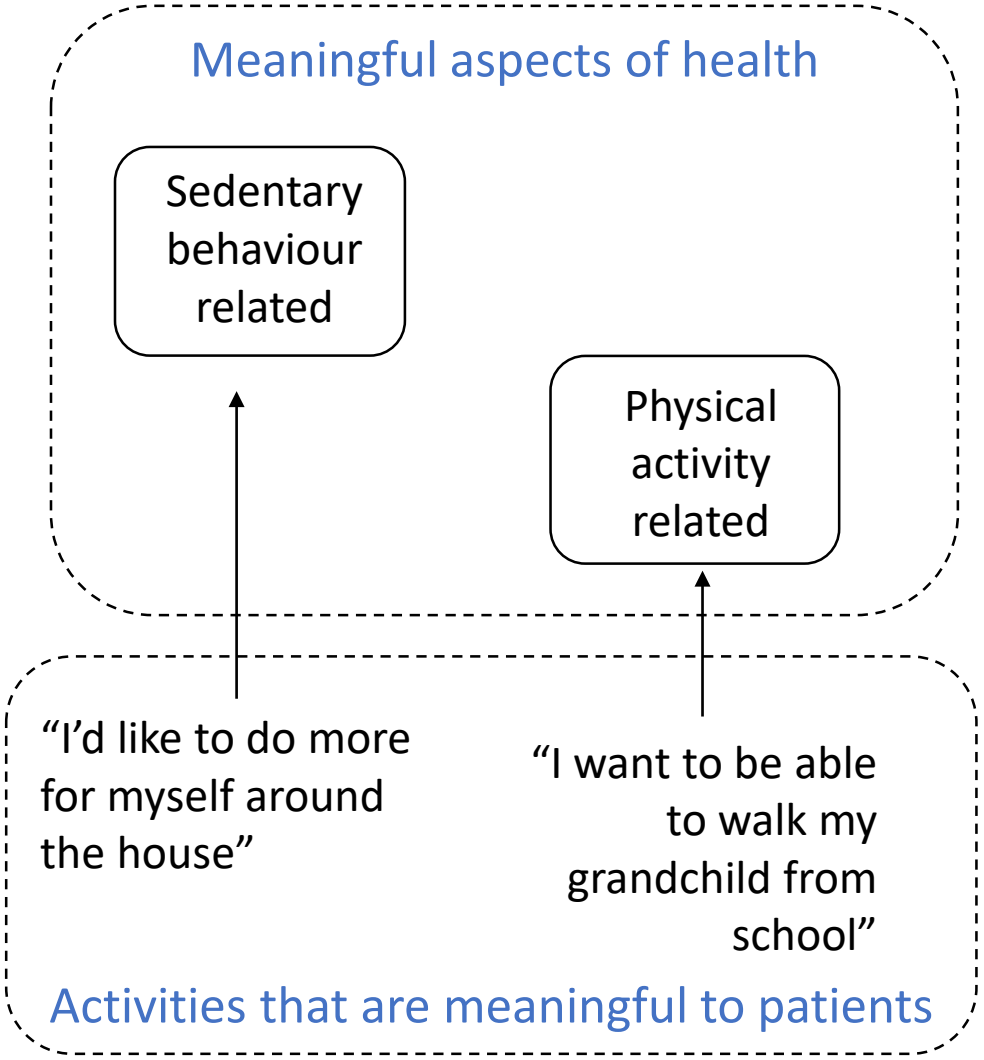
Acknowledgments

- Marc K Walton, Janssen Research and Development

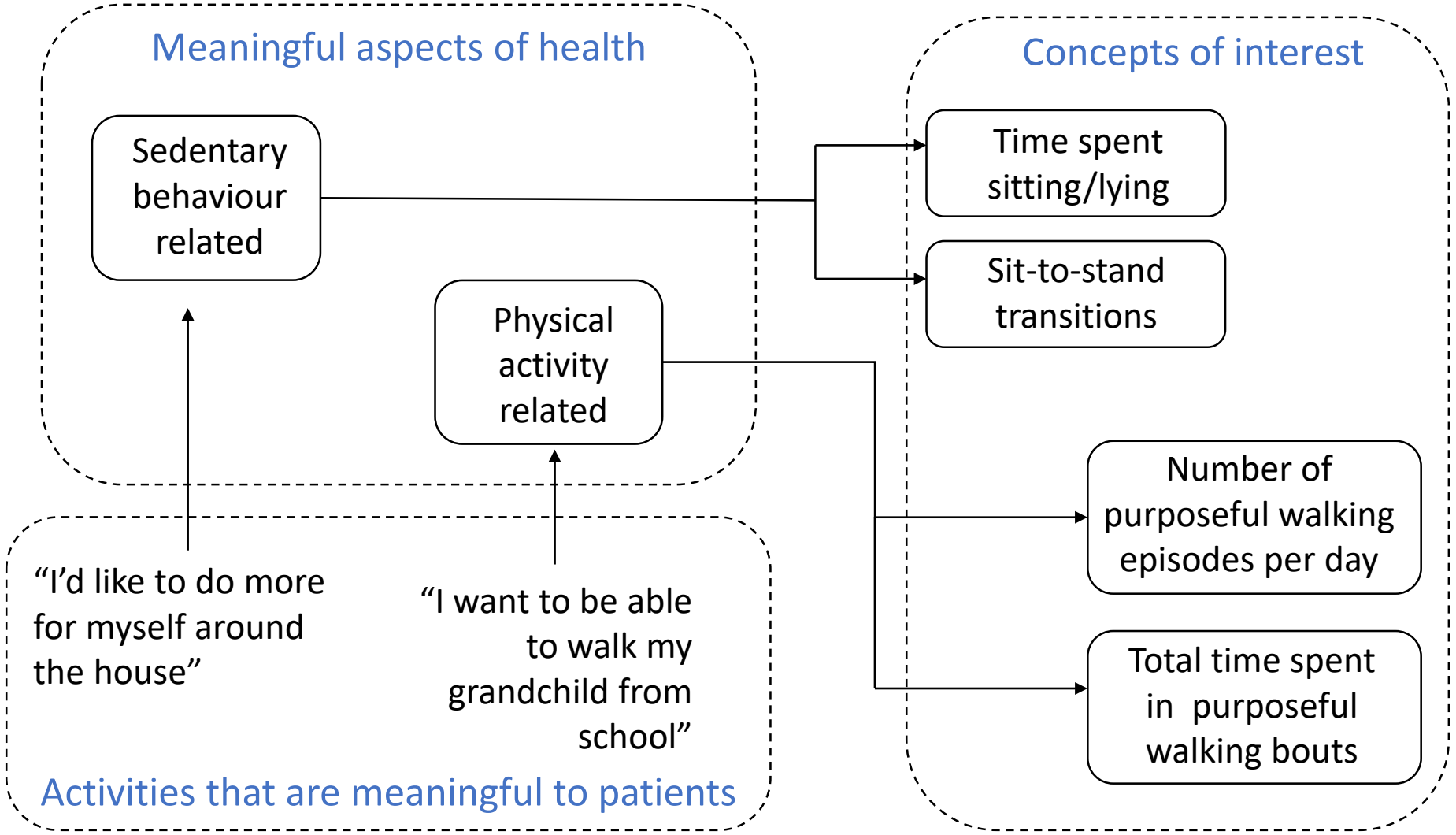
- DIA Study Endpoints Community – Working group on wearables

- Walton MK *et al.* Clinical Outcome Assessments: Conceptual Foundation—Report of the ISPOR Clinical Outcomes Assessment – Emerging Good Practices for Outcomes Research Task Force. *Value Health* 2015; **18**: 741-752.

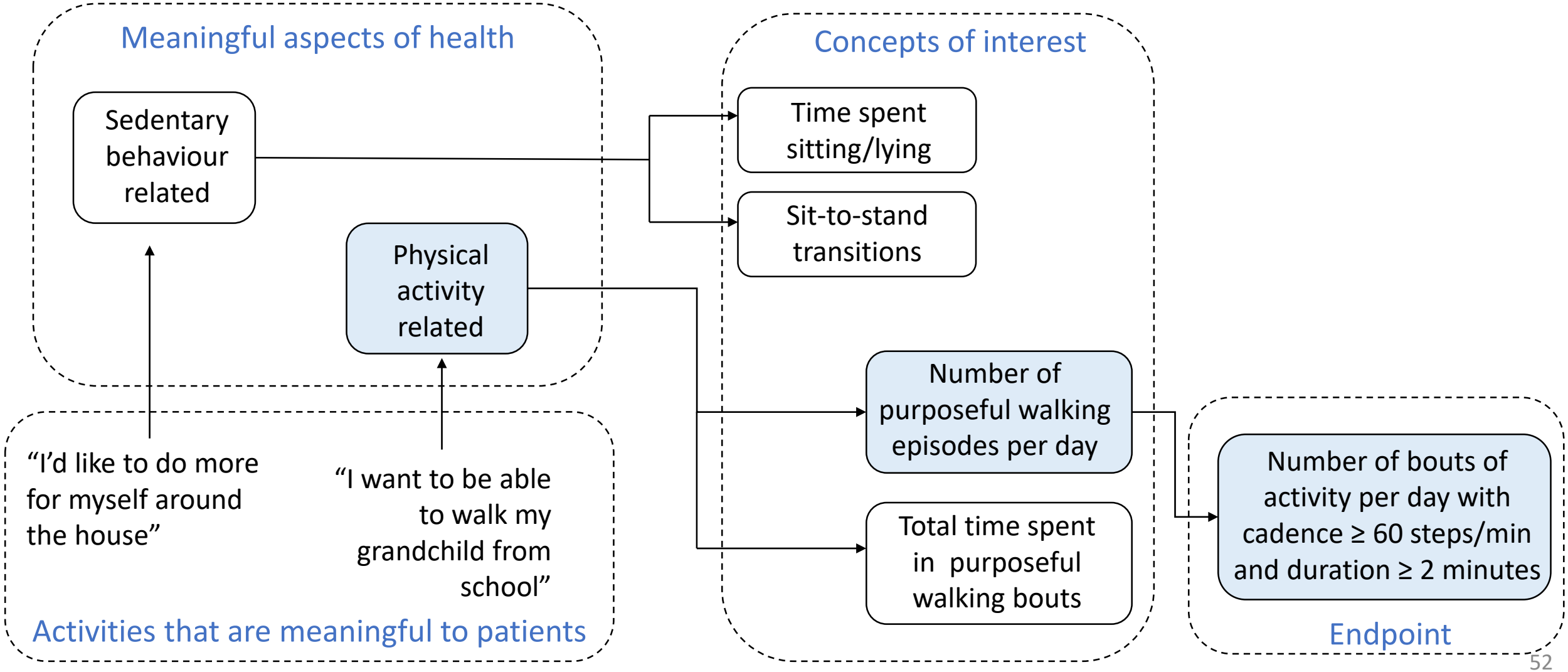
Measuring meaningful aspects of health



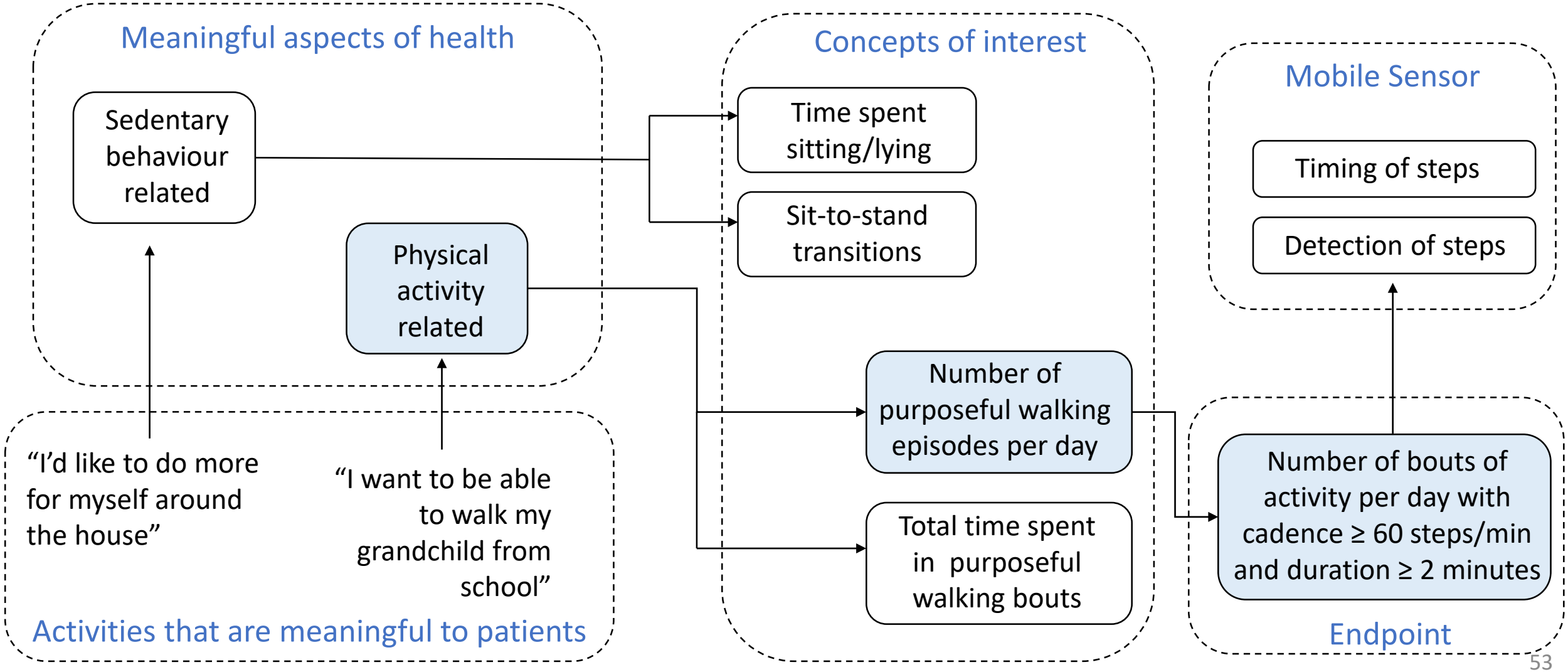
Measuring meaningful aspects of health



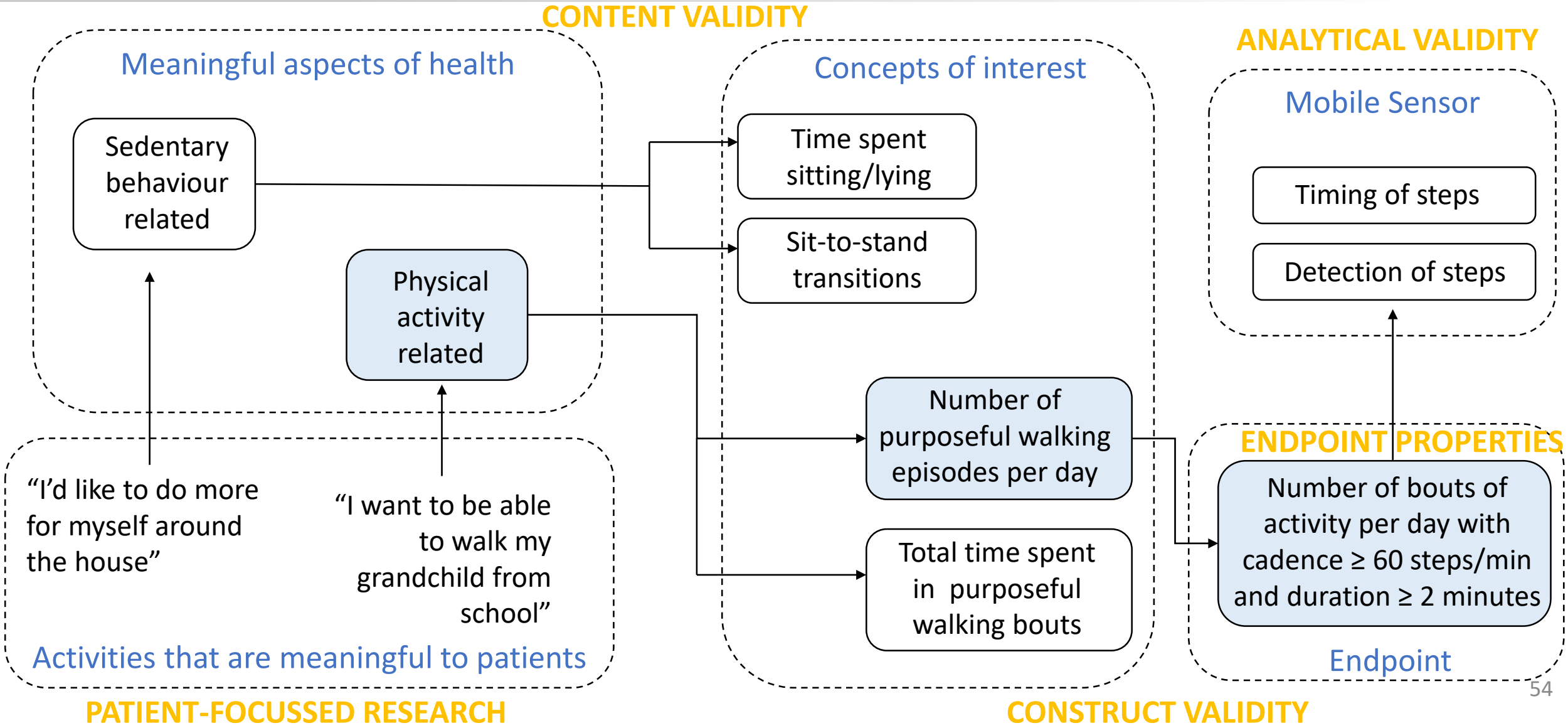
Measuring meaningful aspects of health



Measuring meaningful aspects of health



Measuring meaningful aspects of health



Panel Discussion and Q&A

Moderator

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

Presenters

- *Chad Gwaltney, PhD* – President, Gwaltney Consulting
- *Maria Mattera, MPH* – Assistant Director, PRO Consortium, C-Path
- *Jeremiah (Jay) Trudeau, PhD* - Director, Patient-Reported Outcomes, Janssen Global Services
- *Wayne Amchin, RAC, MIA, MPA* – Senior Consumer Safety Officer, Division of Cardiovascular and Renal Products, OND, CDER, FDA
- *Bill Byrom, PhD* – Vice President of Product Strategy and Innovation, CRF Bracket

Panelist

- *Ebony Dashiell-Aje, PhD* – Reviewer, COA Staff, OND, CDER, FDA