

Physical Activity Measurement Reactivity among Midlife Adults with Elevated Risk for Cardiovascular Disease: Protocol for Coordinated Analyses Across Six Studies

Kiri Baga, Gabrielle M. Salvatore, Iris Bercovitz, Amanda L. Folk, Ria Singh, Laura M. König, Meghan L. Butryn, Jacqueline A. Mogle, Dani Arigo

Submitted to: JMIR Research Protocols
on: October 11, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5
Supplementary Files..... 26
 Figures 27
 Figure 1..... 28
Multimedia Appendixes 29
 Multimedia Appendix 1..... 30



Physical Activity Measurement Reactivity among Midlife Adults with Elevated Risk for Cardiovascular Disease: Protocol for Coordinated Analyses Across Six Studies

Kiri Baga¹ MA; Gabrielle M. Salvatore¹ PhD; Iris Bercovitz¹ MA; Amanda L. Folk¹ PhD; Ria Singh¹; Laura M. König² PhD; Meghan L. Butryn³ PhD; Jacqueline A. Mogle⁴ PhD; Dani Arigo^{1, 5, 6} PhD

¹ Department of Psychology Rowan University Glassboro US

² Faculty of Psychology University of Vienna Vienna AT

³ Weight, Eating, and Lifestyle Science (WELL) Center Department of Psychological and Brain Sciences Drexel University Philadelphia US

⁴ Department of Psychology Clemson University Clemson US

⁵ Department of Family Medicine Rowan-Virtua School of Osteopathic Medicine Stratford US

⁶ Department of Biomedical Sciences Cooper Medical School of Rowan University Camden US

Corresponding Author:

Dani Arigo PhD

Department of Psychology
Rowan University
201 Mullica Hill Road
Robinson 116G
Glassboro
US

Abstract

Background: Cardiovascular disease (CVD) is highly prevalent in the U.S. and adults ages 40-60 with specific health conditions are at particularly elevated risk. Physical activity (PA) is a key cardioprotective behavior and many interventions exist to promote PA in this group. Importantly, effective promotion requires accurate assessment of PA behavior to correctly evaluate intervention effects. As PA is variable and often estimated by averaging across multiple days, a potential threat to accurate assessment is measurement reactivity, or an atypical increase in PA behavior at the start of measurement periods that may bias conclusions. Evidence for PA measurement reactivity is equivocal, though concern has resulted in recommendations to add or drop PA measurement days from inclusion, which may introduce undue burden on participants. At present, the extent of PA measurement reactivity and the behaviors most likely to be affected (e.g., steps vs. minutes of exercise) are unclear. In addition, participant characteristics such as gender and the study expectations (e.g., intervention vs. observation only) may contribute to differences in these patterns, though the extent of these differences is unknown.

Objective: An improved understanding of the extent of PA measurement reactivity and potential moderators among adults in midlife with CVD risk factors may inform decisions about whether and how to account for PA measurement reactivity, and in what contexts.

Methods: To achieve this goal, we will conduct a coordinated multilevel analysis across 6 studies among adults ages 40-60 who have >1 risk factors for CVD (total N = 1844; 13,434 total days of PA measurement with 6-7 days per person). The aims of this project are (1) to estimate the extent of PA measurement reactivity in each study and its influence on estimates (activity units, steps per day, minutes of activity in different intensities), (2) to examine individual difference characteristics as moderators of these patterns (e.g., gender, BMI, PA motivation), and (3) to examine differences in these patterns by study and measurement characteristics (e.g., PA monitoring device, location of wear, observation only vs. intervention baseline). Changes in PA behavior across the measurement period will be examined at the day level, using 2-level multilevel models (days nested within persons) and cross-level interactions (for moderation effects).

Results: Data to be included are from publicly available datasets (observation only: NHANES and MIDUS) and baseline weeks of observation from behavioral weight loss clinical trials (K = 4), all collected in the U.S. Dataset acquisition is complete and data cleaning in preparation for analysis is underway.

Conclusions: Results from this coordinated analysis project will provide the first large-scale estimation of the extent of PA measurement reactivity in an at-risk group. Findings will inform best practices for mitigating potential measurement reactivity in multi-day assessments of PA behavior.

(JMIR Preprints 11/10/2024:67438)

DOI: <https://doi.org/10.2196/preprints.67438>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org>

No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Physical Activity Measurement Reactivity among Midlife Adults with Elevated Risk for Cardiovascular Disease: Protocol for Coordinated Analyses Across Six Studies

Kiri Baga, M.A.,¹ Gabrielle M. Salvatore, Ph.D.,¹ Iris Bercovitz, M.A.,¹ Amanda L. Folk, Ph.D.,¹ Ria Singh,¹ Laura M. König, Ph.D.,² Meghan L. Butryn, Ph.D.,³ Jacqueline A. Mogle, Ph.D.,⁴ & Danielle Arigo, Ph.D.^{1,5}

¹Department of Psychology, Rowan University; ²Faculty of Psychology, University of Vienna; ³Department of Psychology, Drexel University; ⁴Clemson University; ⁵Department of Family Medicine, Rowan-Virtua School of Osteopathic Medicine

Abstract

Background: Cardiovascular disease (CVD) remains the leading cause of death in the U.S. and adults ages 40-60 with specific health conditions are at particularly elevated risk for developing CVD. Physical activity (PA) is a key cardioprotective behavior and many interventions exist to promote PA in this group. Effective promotion requires accurate assessment of PA behavior; as PA is often estimated by averaging across multiple days, a threat to accurate assessment is *measurement reactivity*, or an atypical increase in PA behavior at the start of measurement periods that may bias conclusions. Evidence for PA measurement reactivity is equivocal, though concern has resulted in recommendations to add or drop PA measurement days from inclusion, which may introduce undue burden on participants. At present, the extent of PA measurement reactivity and the behaviors most likely to be affected (e.g., steps vs. minutes of exercise) among those at risk for CVD are unclear, as are participant characteristics such as gender and study expectations (e.g., intervention vs. observation only) that may contribute to differences in these patterns.

Objectives: To improve on the current understanding of the extent of PA measurement reactivity and potential moderators among U.S. adults ages 40-60 with CVD risk factors.

Methods: To achieve this goal, we will conduct coordinated multilevel analyses across 6 studies. Data are from nationally representative, publicly available datasets (observation only: $K=2$) and baseline weeks of observation from behavioral weight loss clinical trials ($K=4$), all collected in the U.S. Publicly available datasets are the National Health and Nutrition Examination Survey (NHANES; 2013-2014) and the Midlife in the United States study (MIDUS; 2004-2009; total $N=1385$), which are available from the Inter-university Consortium for Political and Social Research website. Behavioral weight loss trials were conducted by the Drexel University Weight Eating and Lifestyle (WELL) Center (2011-2023; total $N=444$), in person or remotely via Zoom. Relevant data

from each study were extracted for adults ages 40-60 who have ≥ 1 risk factors for CVD (total $N=1832$; 11,707 total days of PA measurement with 6-7 days per person). Changes in PA behavior across the measurement period will be examined at the day level, using 2-level multilevel models (days nested within persons) and cross-level interactions (for moderation effects).

Results: This project was funded in August 2022 and received supplementary funding in September 2023. Dataset acquisition and data cleaning were completed in October 2024. Analyses are expected to be completed in February 2025 and findings are anticipated to be shared in June 2025.

Conclusions: Results from this coordinated analysis project will provide the first large-scale estimation of the extent of PA measurement reactivity in an at-risk group. Findings will inform best practices for mitigating potential measurement reactivity in multi-day assessments of PA behavior.

Keywords: physical activity; measurement reactivity; gender difference; cardiovascular risk; intensive assessment

Introduction

Cardiovascular disease (CVD) remains the leading cause of death in the U.S., and risk for CVD meaningfully increases during midlife (ages 40-60) [1,2]. Increased risk for CVD during midlife is due to universal aging processes and associated conditions such as obesity and hypertension, as well as to sex-specific metabolic and hormonal changes (e.g., the onset of menopause), which result in distinct risk patterns for women and men [2-6]. Engaging in regular physical activity (PA) during midlife can significantly reduce risk for CVD for both women and men [7,8]. Unfortunately, low PA engagement among midlife adults is common, which exacerbates their CVD risk and contributes to their substantial healthcare costs [4,9]. This problem persists despite decades of efforts to promote PA that include high participation rates among midlife adults [10,11], and when PA promotion is effective, the benefits of participation are short-lived [12]. This applies to the many PA interventions that are specifically tailored for midlife adults, including those that are gender-specific (i.e., enroll only men or women; focus on sex- or gender-specific content such as PA in the context of gender roles, or effects of PA on menopause symptoms or risk for prostate cancer) [13-18]. Evidence consistently shows that very few midlife adults – particularly those with CVD risk conditions such as obesity, hypertension, or type 2 diabetes – achieve PA levels that would confer significant protection against CVD [19-21]. Thus, additional work is sorely needed to effectively promote PA to reduce CVD risk among midlife adults.

Critically, however, our understanding of current PA engagement and the benefits of PA for reducing CVD risk among midlife adults relies on accurate PA measurements in this at-risk population. Assessment of PA using research-grade or commercially available devices (e.g., Fitbit) provides more accurate estimates of PA engagement than methods such as self-report [22]. To ensure accurate estimates by accounting for normal daily variability in PA engagement [23], monitoring periods typically last 7 days, and observed PA behavior (e.g., steps per day, minutes of moderate-to-vigorous-intensity activity) is averaged across these days to generate an overall estimate of PA engagement [24]. Evidence across the multidisciplinary field of PA promotion highlights the potential for bias in the assessment of PA using monitoring technology; one source of this bias is *measurement reactivity*, or change in PA behavior due to the introduction of its measurement [17,21,25-26]. Introducing PA monitoring via external device during participation in a research study is thought to increase the salience of both immediate PA engagement and the notion

that someone else is watching [27-28]. This increased awareness can lead to significantly more PA engagement early in the monitoring period than is typical or sustainable. Specifically, PA measurement reactivity typically manifests as greater PA at the start of observation periods, relative to subsequent days (e.g., negative linear slope across days, or higher on Day 1 or Days 1-2) [25]. When estimates are calculated as averages across days, initial elevation could skew these estimates to show greater engagement than is typical or sustainable [25]. The resulting inflation of PA estimates could mask the effects of PA interventions, as pre-intervention starting points could be overestimated [17] (among other confounds), leading to incorrect conclusions about PA engagement among at-risk groups such as adults in midlife.

To fully understand the value of promoting PA to reduce CVD risk among adults in midlife with existing risk factors, it is critical to ensure the accuracy of PA measurements in this population. As noted, if measurement reactivity effects are substantial, the resulting biased estimates of PA in at-risk groups such as midlife adults could lead to incorrect conclusions and next steps for PA promotion. PA measurement reactivity has been observed in a subset of studies to date [25,29-31] and some PA researchers consider reactivity a critical source of bias that warrants increased attention and mitigating actions, such as requiring extra days of observation to get used to the PA monitor or removing the first 1-2 days of PA observation from analyses [30]. The evidence is equivocal, however, with some studies showing no reactivity patterns or patterns that are not clinically significant (and thus, would not meaningfully affect overall PA estimates) [25,28]. As a result, there is considerable debate about the need for such drastic measures, or even the need for ongoing attention to this phenomenon. PA recordings on these days represent actual PA engagement, even if the level of engagement is not “typical,” and adding assessment days increases participant burden [32]. Consequently, the extent of PA measurement reactivity and best practices for handling PA data to address the potential for it to affect estimates have yet to be identified.

Existing evidence is also limited in three key ways. First, studies of PA measurement reactivity have often focused on children or healthy, younger adults. Because these populations generally are more active than midlife adults with elevated risk for CVD (or similar at-risk groups) [19,33], they may respond differently to the introduction of PA monitoring. Second, there is little consistency between existing studies with respect to definitions of reactivity or which PA outcomes are affected; outcomes of interest have included steps per day, light PA, moderate-to-vigorous PA (MVPA), and sedentary time [29,30,33], which have distinct implications for CVD risk that differ by gender [34-35]. In addition, some work has defined reactivity as an overall linear decrease in PA across days of observation [25,36], whereas others have examined direct comparisons between Day 1 and subsequent days (Days 2, 3, 4, or combinations) [29,37]. Third, there has been little attention to the contexts where PA measurement reactivity is a meaningful problem, leading to inconsistent and often drastic recommendations for addressing it. As noted, these include adding observation days, which increases participant burden, and removing the first 1-2 days of observation, which reduces the precision of estimates (or both of these steps) [25,27,30,38]. Although there is some convergence in PA estimates between commercially available wearable devices and research-grade monitors [39], differences in location of wear (e.g., wrist vs. hip) and the extent to which a device offers PA feedback (e.g., via digital display of accumulated steps per day in real time) could influence reactivity, by differentially increasing the salience of PA [31]. Assessment of PA prior to the start of an intervention (vs. in studies that use observation alone) may also affect reactivity via differences in participants' anticipation of future increases in their PA [30].

In addition, individual differences in mental health symptoms, physical health characteristics, and psychosocial processes are known determinants of PA engagement and thus, may influence reactivity patterns [39]. For example, depressive symptoms, number of cardiovascular risk factors, and body

mass index (BMI) are negatively associated with PA engagement [40-41] and as a result, may limit reactivity. In contrast, tendencies toward self-evaluation relative to others and increased motivation for health behaviors may have the opposite effects [40,42]. Specifically, women and those who have stronger tendencies to make and value social comparisons (i.e., self-evaluations of one's traits or behaviors compared to others [43]) may show greater measurement reactivity responses, as these individuals may be particularly interested in managing others' perceptions of them (such as researchers) [43-46]. Higher (vs. lower) levels of PA motivation and more (vs. fewer) past attempts at PA behavior change may also be associated with stronger reactivity responses, as these may indicate greater focus on PA overall [47]. Greater attention to context, rather than treating PA measurement reactivity as universally problematic versus not (as it is currently), could identify optimal and tailored targets for mitigating reactivity in future studies.

Thus, the role of PA measurement reactivity in PA promotion for at-risk groups remains unclear, and no existing study has examined the extent of reactivity based on research design, study procedures, and individual participant characteristics. **For the present study, we will capitalize on the availability of existing data resources with a multilevel coordinated analysis across 6 studies (including 2 nationally representative observational studies and 4 intervention trials), to achieve the following aims:**

- 1) **To characterize midlife adults' PA measurement reactivity in each dataset.** We will determine whether daily PA engagement meaningfully changes over 6-7 days of observation, across available PA outcomes. Available data include Monitor-Independent Measurement Summary (MIMS) units, activity counts, steps per day, minutes of light activity per day, and minutes of moderate-to-vigorous- intensity activity (MVPA) per day.
- 2) **To determine whether the presence or extent of measurement reactivity differ based on demographic, medical, or psychological characteristics.** We hypothesize that reactivity will be weaker among adults with higher (vs. lower) BMIs, numbers of CVD risk factors, and depressive symptoms. Conversely, we hypothesize that reactivity effects will be stronger among women, adults with higher (vs. lower) PA motivation and social comparison tendencies, and those with more previous attempts to increase PA or lose weight. We will further examine whether gender moderates the effects of other individual difference characteristics on PA engagement over 6-7 days, by including the interaction effect between gender and other predictors.
- 3) **To determine whether the presence or extent of measurement reactivity differ based on study characteristics.** We hypothesize that reactivity effects will be stronger in studies that use (a) observation only (as there is no expectation of later improvement), (b) commercially available devices (which tend to be visible throughout the day), (c) wrist wear (based on both visibility and sensitivity to movement [31,48]), and (d) devices that offer feedback in real time (as they provide additional information that can prompt a behavioral response [49]). We will further examine whether gender moderates the effects of study characteristics on PA engagement over 6-7 days, by including the interaction effect between gender and other predictors.

Methods

Overview

Using a coordinated analysis approach [50], we will model differences in PA engagement across days of observation using data from each of 6 studies. These include 2 observational studies that have publicly available datasets and 4 clinical trials datasets. All data were collected between January 2005 and January 2023. Procedures included 6-7 days of intensive PA assessment using a

research-grade and/or commercially available device. Analysis of available datasets leverages existing resources to address questions beyond those originally intended; this limits cost and eliminates additional participant burden while offering insights into important phenomena [51]. This approach is uniquely suited for the present study, as the concern about PA measurement reactivity is that it appears any time PA monitoring devices are introduced. If this concern is warranted, evidence of measurement reactivity should be observed in studies that were not specifically designed to test this phenomenon (e.g., studies for which PA was assessed for other purposes and data are already available). The datasets used for this set of secondary analyses were selected for their accessibility to the research team, representation of a large number and range of individuals in the target population, and heterogeneity in study designs and monitoring PA devices used. Together, these datasets will enable examination of reactivity patterns following the introduction of PA measurement devices in multiple people and contexts.

Observational Studies

These include samples from the National Health and Nutrition Examination Survey (NHANES) [52] and the Midlife in the United States (MIDUS) [1] Study (combined $n=1385$). NHANES is a national, longitudinal program of research focused on understanding various aspects of health among adults and children in the U.S., and data are publicly available. We will use data from the most recent wave of collection that included PA monitoring (2013-2014; 7 days of assessment with the ActiGraph GT3X). MIDUS is a longitudinal investigation of life circumstances and health outcomes among adults aged 25-75. The current study uses data from the MIDUS Biomarker Project (2004-2009), which included PA monitoring (6 days of assessment with the Actiwatch-64).

Randomized clinical trials

These studies were tests of improvements to standard behavioral weight loss treatment ($K=4$, combined $n=444$). Each study employed a 7-day pre-treatment observation window, which will be used to examine PA measurement reactivity. Project ENACT (NCT01858714) tested the effects of enhanced focus on the food environment and the use of acceptance and commitment skills to support behavior change on long-term weight loss [53]. Collection of the data to be used in the planned analyses began in 2011 and finished in 2013. Project IMPACT (NCT02363 010) tested an increased emphasis on PA and acceptance and commitment skills to support behavior change on long-term weight loss [54]. Collection of the data to be used in the planned analyses began in 2014 and finished in 2016. The last two studies examined the added benefits of sharing PA self-monitoring data, either with coaches (Project FitLink Pilot, 2018-2019; NCT03337139 [55]) or with coaches, a designated member of the participant's existing social network (friend or family member), and/or other participants (Project FitLink Full, 2021-2023; NCT05180448 [56]). ENACT and IMPACT used the same research-grade accelerometer (Actigraph GT3X), worn on the hip. FitLink Full used a commercially available wearable PA monitor (Fitbit Inspire 2), worn on the wrist, and FitLink Pilot used both the research-grade hip-worn accelerometer and the commercially available wrist-worn wearable described above (see Table 1). The commercial device was worn from the baseline period through treatment; the research-grade device was added at a 3-month assessment (i.e., at randomization to condition after a uniform treatment period). ENACT, IMPACT, and FitLink Pilot were conducted in person in a large city in the northeastern U.S., whereas FitLink Full was conducted remotely using national recruitment and enrollment.

Participants

Eligible participants are adults in midlife (ages 40-60) with ≥ 1 risk factors for CVD who completed the relevant PA assessment period (see below for additional details about inclusion in data analyses). CVD risk factors include prediabetes, type 2 diabetes, hypertension, high

cholesterol, obesity, and current smoker. As noted, observational studies ($n=1385$) recruited nationally in the U.S.; clinical trials ($n=444$) were behavioral weight loss studies completed in the northeastern U.S., one of which was conducted remotely. Participant demographics are listed in Table 2.

Procedures

Datasets were chosen for the availability of consecutive days of PA measurement captured immediately following the introduction of commercial and research-grade accelerometers. PA data were recorded using research-grade or commercially available devices (see Table 1). Participant demographics and individual difference characteristics were collected using questionnaires or structured interviews. Specific details about these measures are described in detail below.

Table 1. *Physical activity measurement by study.*

Study Type	Study	Device	Duration	Outcomes
Observational Studies	NHANES ($n=1208$)	ActiGraph GT3X (wrist-worn)	7 days	Monitor Movement (MIMS) units Independent Summary
	MIDUS ($n=177$)	Actiwatch-64 (wrist-worn)	6 days	Total activity count units
Randomized Clinical Trials (Intervention Programs)	ENACT ($n=143$)	ActiGraph GT3X+ (hip-worn)	7 days	Steps per day, minutes of light activity, minutes of moderate-to-vigorous activity
	IMPACT ($n=140$)	ActiGraph GT3X+ (hip-worn)	7 days	Steps per day, minutes of moderate-to-vigorous activity
	Fitlink Pilot ($n=31$)	Fitbit ActiGraph GT3X (hip-worn; week 13)	7 days	Steps per day, minutes of moderate-to-vigorous activity
	FitLink Full ($n=130$)	Fitbit	7 days	Steps per day, minutes of moderate-to-vigorous activity

Measures

Demographics

Participants' basic demographics were collected via self-report. These data included age, gender, racial/ethnic identity, marital status, income, and education (see Table 2). NHANES collected these data during structured, in-person interviews with participants. The remaining studies captured these data using electronic or paper questionnaires.

Physical Activity

Included studies captured 6-7 days of PA data using different PA monitoring devices and will enable examination of device-specific effects. ActiGraph devices were used for observational studies and 3 (of 4) clinical trials. These devices were wrist-worn and hip-worn ActiGraph models for observational studies and clinical trials, respectively; remaining clinical trials used commercially available Fitbit devices (see Table 1). PA outcomes were daily PA summary metrics, steps per day, or minutes of activity at different intensities. Observational studies used Monitor Independent Movement Summary units (MIMS [57]) and total activity counts; clinical trials datasets provided steps per day and minutes of light and/or MVPA per day (also detailed in Table 1). Steps per day captured using Fitbit at baseline and both a Fitbit and ActiGraph at 3 months into treatment during the FitLink Pilot will allow for comparison of potential recurrence of reactivity patterns after the onset of behavioral weight loss treatment.

Table 2. *Participant demographics across six studies (mean $\pm$ sd or n (%)).*

	NHANES ^a (n=1208)	MIDUS (n=177)	ENACT (n=143)	IMPACT (n=140)	FitLink Pilot (n=31)	FitLink Full (n=130)
Age	49.9 $\pm$ 6.0	50.9 $\pm$ 5.8	52.3 $\pm$ 5.6	52.9 $\pm$ 5.0	50.9 $\pm$ 6.1	51.5 $\pm$ 5.7
Gender						
Women	628 (52.0)	106 (59.3)	117 (81.8)	116 (82.9)	28 (90.3)	109 (83.8)
Men	580 (48.0)	72 (40.7)	26 (18.2)	24 (17.1)	3 (9.7)	21 (16.2)
BMI (kg/m ²)	31.2 $\pm$ 7.4	32.0 $\pm$ 7.2	36.3 $\pm$ 4.5	35.5 $\pm$ 4.4	36.3 $\pm$ 4.5	36.5 $\pm$ 4.8
BMI Category						
<18.5 kg/m ²	9 (0.8)	1 (0.6)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
18.5 – 25 kg/m ²	224 (18.6)	26 (14.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
25 – 30 kg/m ²	321 (26.6)	43 (24.3)	7 (4.9)	6 (4.4)	0 (0.0)	4 (3.1)
>30 kg/m ²	653 (54.1)	107 (60.4)	136 (95.1)	131 (95.6)	31 (100.0)	126 (96.9)
Race ^a						
AI/NA ^b	-	0 (0.0)	1 (0.7)	0 (0.0)	1 (3.2)	0 (0.0)
Asian	115 (9.8)	2 (1.1)	1 (0.7)	1 (0.7)	0 (0.0)	3 (2.3)
NH or OPI ^b	-	0 (0.0)	1 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)
Black/AA ^b	273 (23.4)	50 (28.4)	51 (35.7)	35 (25.0)	12 (38.7)	17 (13.1)
White	486 (41.6)	103 (58.5)	84 (58.7)	98 (70.0)	14 (45.2)	107 (82.3)
Other/Mixed Race	32 (2.7)	21 (11.9)	5 (3.5)	6 (4.3)	4 (12.9)	3 (2.3)
Ethnicity ^a						
Hispanic or Latino	273 (22.6)	0 (0.0)	8 (5.6)	5 (3.6)	2 (6.4)	10 (7.7)
Not Hispanic or Latino/a	935 (77.4)	177 (100.0)	134 (94.4)	135 (96.4)	29 (93.6)	120 (92.3)
Marital Status						
Married	721 (59.2)	100 (56.5)	86 (59.3)	91 (65.0)	15 (48.4)	90 (68.7)
Widowed	28 (2.3)	4 (2.3)	2 (1.4)	2 (1.4)	1 (3.2)	1 (0.8)
Divorced	203 (16.6)	32 (18.1)	22 (15.2)	21 (15.0)	4 (12.9)	6 (6.2)
Separated	52 (4.3)	9 (5.1)	3 (2.1)	5 (3.6)	3 (9.7)	-
Never married	151 (12.4)	32 (18.1)	32 (22.1)	21 (15.0)	8 (25.8)	-
Single	-	-	-	-	-	20 (15.3)
Cohabiting	63 (5.2)	-	-	-	-	9 (6.9)
Not cohabiting	-	13 (16.9)	-	-	-	2 (1.5)

Income ^a						
\$0 to \$25k	294 (25.5)	63 (36.6)	9 (6.4)	4 (2.9)	1 (3.2)	-
\$25k - \$50k	210 (18.2)	53 (30.8)	21 (14.9)	13 (9.5)	4 (12.9)	-
\$45k - \$55k ^a	92 (7.9)	-	-	-	-	-
\$50k - \$75k	122 (10.6)	27 (15.7)	31 (22.0)	15 (11.0)	7 (22.6)	-
\$75k - \$100k	113 (9.8)	14 (8.1)	24 (17.0)	23 (16.8)	6 (19.4)	-
>\$100k	251 (21.8)	8 (4.6)	13 (9.2)	20 (14.6)	2 (6.5)	-
\$125k - \$150k	-	2 (1.2)	18 (12.8)	17 (12.4)	6 (19.4)	-
\$150k - \$175k	-	2 (1.2)	8 (5.7)	18 (13.1)	0 (0.0)	-
\$175k - \$200k	-	0 (0.0)	5 (3.6)	13 (9.5)	2 (6.4)	-
>\$200k	-	3 (1.7)	12 (8.5)	14 (10.2)	3 (9.7)	-
Education						
Less than 9 th grade	80 (6.8)	2 (1.1)	0 (0.0)	0 (0.0)	0 (0.0)	-
Partial high school	156 (13.4)	10 (5.6)	1 (0.7)	0 (0.0)	0 (0.0)	-
High School or GED	280 (24.0)	35 (19.8)	13 (9.1)	5 (4.3)	1 (3.3)	-
Associate's degree, technical degree, or partial college	373 (31.9)	52 (43.5)	22 (15.4)	19 (16.4)	4 (12.9)	-
Bachelor's degree	279 (23.9)	44 (24.9)	52 (36.4)	41 (35.3)	11 (35.5)	-
Graduate or professional degree	-	34 (19.2)	55 (38.5)	51 (44.0)	15 (48.4)	-

^aNHANES specific notes: Race and ethnicity questions were combined for NHANES; Income was based on household, not individual, and had different cutoff values. ^bRace categories: AI/NA – American Indian/Native Alaskan; NH or OPI – Native Hawaiian or Other Pacific Islander; AA – African American

Individual Difference Measures (see Table 3)

Table 3. Measures of individual difference characteristics by study.

Characteristic	NHANES	MIDUS	ENACT	IMPACT	Fitlink Pilot	FitLink Full
BMI	Measured	Measured	Measured	Measured	Measured	Measured
CVD Risk	Individual questions (diagnosis)	Individual questions (diagnosis)	WALI BMI	WALI BMI	WALI BMI	WALI BMI
Depressive Symptoms	PHQ-9	CES-D	BDI-II	BDI-II CES-D	WALI	BDI-II
PA Motivation					TSRQ	TSRQ
Weight loss attempts	Self-reported (past year)		WALI	WALI	WALI	WALI
Social		Compariso	INCOM			INCOM

comparison		n of CVD risk				
------------	--	---------------	--	--	--	--

CVD risk conditions include prediabetes or type 2 diabetes, hypertension, high cholesterol, obesity, or current smoker. Participants were prompted to report whether they had a physician diagnosis of prediabetes, type 2 diabetes, hypertension, and/or high cholesterol. Body mass index (BMI) ≥ 30 kg/m² will be used to indicate obesity and will be calculated based on measured height and weight. Table 4 summarizes the risk factors assessed in each study. Participants with pre-existing CVD will be excluded from analyses, including those who self-report a history of heart disease, heart failure, stroke, or heart attack. We will calculate the proportion of CVD risk factors endorsed out of the risk factors assessed, to determine an overall risk percentage.

Table 4. *CVD risk conditions assessed by study.*

CVD risk condition	NHANE S	MIDUS	ENACT	IMPACT	FitLink Pilot	FitLink Full
Hypertension	•	•	•	•	•	•
High cholesterol	•					
Prediabetes	•					
Type 2 Diabetes	•	•	•	•	•	•
Obesity (BMI ≥ 30 kg/m ²)	•	•	•	•	•	•
Smoker	•	•	•	•	•	

Depressive symptoms were measured in each study using validated self-report measures, including the Patient Health Questionnaire (PHQ-9 [58]), the Center for Epidemiologic Studies Depression Scale (CES-D [59]), and the Beck Depression Inventory (BDI [60]). Total scores for these measures range from 0-27, 0-60, and 0-63, respectively, with higher scores indicating more severe depressive symptoms. Depressive symptoms were captured for the FitLink Pilot using the Weight and Lifestyle Inventory (WALI [61]), which prompts participants to report experience of depressed mood or anhedonia during the past month.

Social comparison will be examined for 3 datasets (i.e., MIDUS, ENACT, FitLink Full). MIDUS captured health-related social comparison experiences; participants were prompted to indicate their perception of their level of risk for a heart attack compared to others (i.e., higher, lower, or the same), and the degree of this difference (e.g., a lot higher, somewhat higher, or only a little higher). Social comparison experiences for clinical trials participants were measured using the Iowa–Netherlands Comparison Orientation Measure (INCOM [43]). This measure assesses respondents' perceptions of their own tendency to make social comparisons, generally and tendencies toward upward (i.e., better off) and downward (i.e., worse off) comparison targets. Higher total scores indicate stronger tendencies toward social comparison.

Motivation to be physically active was assessed for 4 datasets (clinical trials; i.e., ENACT, IMPACT, Fitlink Pilot, Fitlink Full) using an adapted version of the Treatment Self-Regulation

Questionnaire (TSRQ [62]). This measure assesses different types of motivation for changing health behaviors by capturing the extent to which respondents endorse particular reasons for changing their health behavior on a scale from 1 (*not at all true*) to 7 (*very true*). Specific TSRQ item scores are summed to calculate summary scores for autonomous motivation, introjected regulation, external regulation, and amotivation, with higher scores indicating greater endorsement of motivation subtypes.

History of weight loss attempts was collected in 5 studies. All 4 clinical trials used the Weight and Lifestyle Inventory (WALI [61]), which asks respondents to report their age at the time of these efforts, weight loss method, and pounds lost. These occurrences were summed to indicate participants' history of weight loss attempts. NHANES collected these data by prompting participants to indicate if they attempted weight loss during the past year, and if so, asking whether they used specific weight loss methods. Participants' responses to these prompts will be summed to indicate the total number of weight loss methods used during the past year.

Ethical considerations

The coordinated, secondary analysis outlined here is approved as exempt (no human subjects enrollment) by the Institutional Review Board at Rowan University and Rowan-Virtua School of Osteopathic Medicine (protocol number PRO-2021-550). **Original data collection for public datasets and clinical trials was approved by the organizations' home institutional review boards and participants received monetary compensation for their time and effort [1, 52-56]. All consent processes involved authorization for secondary analyses. Datasets were anonymized or de-identified prior to analyses.**

General Analysis Plan

Our primary analyses will examine how device-assessed PA behavior varies within-persons across days of participation, using multilevel models (MLM). Each dataset includes summary indicators of PA on a given day and days are nested within persons, creating dependence best accounted for by two-level MLMs. Primary outcomes of interest will be MIMS units, activity counts, steps, minutes of light activity, and minutes of MVPA, as available in each dataset. As noted, data from participants who met demographic and medical criteria were included in the present analyses. We also set minimum PA monitoring device wear time to 10 hours of wake time per day, and days with <10 hours of valid wear time were excluded from analyses. **The original, reduced, and analytical sample sizes for each study based on these criteria are provided in Figure 1.**

Coordinated Analysis Approach

As we have multiple datasets that share our outcomes and moderators of interest, we will use coordinated analysis to arrive at conclusions [50]. **In coordinated analysis, models are parameterized and fit to datasets and statistical results from specific datasets are compared and contrasted, to efficiently replicate substantive conclusions [63-64].** Given the differences in study design, we will fit equivalent models separately across datasets and standardized estimates can be generated to determine the consistency of effects, with respect to direction and relative size. As in our previous work, we will use semipartial correlation coefficients as standardized effect size estimates [65-68]. Fitting separate models also allows for exploration of dataset-specific covariates and associations, to address substantive questions relevant to individual datasets.

Aim 1. To address our first aim, we will fit models that examine whether daily parameters of PA behavior at level 1 change within a person as a function of day of measurement, to describe the pattern of measurement reactivity. For these models, we will examine day of measurement as a

continuous predictor of PA using both linear and quadratic trends. Further, we have the opportunity to consider the day of measurement as a categorical predictor to identify whether a particular day of measurement shows greater or lower activity counts relative to adjacent days. As in previous work, we will document patterns where Days 1 and Days 1-2 meaningfully differ from other days of observation as measurement reactivity [29,37,65,69].

Aim 2. We will then add moderators of change across days to address our second aim: to understand whether patterns of measurement reactivity differ based on BMI, percentage of CVD risk factors assessed, depressive symptoms, and gender. We will also examine whether patterns of measurement reactivity depend on motivation to engage in PA, social comparison responses, and previous attempts to lose weight. All moderators will be entered at level 2 as between-person predictors and the cross-level interaction with day of measurement included, to assess for differences in patterns of measurement reactivity. We will also examine whether any reactivity that does emerge differs between genders. To do this, we will incorporate the 3-way interaction among gender, day of measurement, and any of the above moderators. Significant interactions will be probed with specific contrasts and visualizations separately for each group for categorical moderators or at ± 1 standard deviation of the mean for continuous moderators.

Aim 3. To address our third aim, we will use standardized estimates to compare the effect of day of measurement on PA summary scores across study designs and device characteristics. Consistent with our follow-up analyses for our second aim, we will incorporate a cross-level interaction with gender and day of measurement to determine whether any effects of study design and device type depend on the gender of the individual being measured.

Results

The datasets described provide 11,707 valid PA observations across 1832 adults in midlife with ≥ 1 risk factors for CVD. As indicated, participant characteristics for each study can be found in Table 2. Dataset acquisition is complete and data cleaning in preparation for analysis is underway. We expect to complete analyses by March 2025 and to make full results available by June 2025.

Discussion

This study will be the first to systematically investigate PA measurement reactivity in a population with elevated CVD risk using a coordinated multilevel analysis. Analyses will be executed on a large, diverse sample of midlife adults with elevated risk for CVD, who are frequent targets of PA promotion and who make up substantial subgroups in broader PA programs [2-7]. A critical advantage of the proposed analytic approach is that days will be treated separately and nested within individuals, to determine the extent of potential effects of reactivity on estimates of averages [22]. This approach will also be applied across multiple research designs, PA monitoring devices, and individual difference characteristics, allowing for identification of potential moderators of PA measurement reactivity responses.

A noteworthy limitation of our approach is that there is no opportunity to compare the same PA outcome across all studies. Only certain variables are publicly available for NHANES and MIDUS (i.e., MIMS units and total activity counts) and these cannot be converted to estimates of steps or minutes of activity per day [70]. Conversely, only the latter are available for clinical trials and these cannot be converted to corresponding units. To address this, we will report and compare standardized effect sizes across outcomes and studies [24]. In addition, a subset of participants in each study did not contribute enough valid PA data to be included, and participants who were included occasionally had days where valid data were missing due to insufficient wear time on that

day (<20% of expected days) [25]. An additional advantage of a multilevel modeling approach is that these models are robust to missing data [71-72]. Thus, estimates and conclusions are unlikely to be affected by the observed level of missingness at the day level [25,28]. At the person level, we will also compare participants who were included in analyses to those who were excluded for having too few days of valid PA data. This will allow us to determine whether these groups differ with respect to demographic or medical characteristics and to what extent this level of missingness might impact the generalizability of our findings.

As noted, there is considerable debate about the need for dramatic steps to address PA measurement reactivity, such as adding and/or removing days of observation or blinding participants to PA feedback from their measurement devices [25,29-31,73]. Limitations notwithstanding, findings from this study will indicate whether such steps are warranted among adults in midlife with elevated CVD risk and under what circumstances (e.g., across participants or study designs vs. for certain subgroups or study designs). Adults in this population are primary targets and recipients of PA assessment research and intervention, as evidence consistently shows low PA in this group despite considerable PA promotion efforts. However, very little PA measurement reactivity work has focused on this group. Drastic measures to minimize reactivity responses may improve PA estimates if reactivity is substantial; if not, such measures are likely to limit the benefits of self-monitoring in an intervention [74], as well as waste resources and increase participant burden for those who are most in need of support. Alternatively, there may be subgroups for whom or research contexts in which PA measurement reactivity is particularly problematic. Focusing attention on low-burden ways to minimize reactivity responses in situations where it is most likely to bias PA estimates would be more cost-effective than universal measures. For example, introducing participants to the concept and the possibility of reactivity, encouraging participants to be aware of it, and emphasizing engagement in normal PA behavior despite research participation or the presence of a monitor may effectively mitigate effects on PA estimates [65,75]. Such an approach, rather than adding days of assessment, would also increase accessibility of research and intervention for midlife adults, who often have busy, unpredictable schedules [76] and who may find additional days of assessment overwhelming (and therefore, decline to participate or withdraw after enrollment [77-78]). Findings from this study will provide high-quality evidence to determine the effect of measurement reactivity on PA estimates in a key population of interest, and thereby, contribute to best practice recommendations for measuring PA in daily life [31,38].

Acknowledgements

The authors would like to thank Natasha DeMeo, M.Sc., for her assistance with data acquisition and management.

Funding

This work was supported by the U.S. National Institutes of Health under NHLBI R03160602 and a competitive administrative supplement from the Office of Research on Women's Health (PI: D. Arigo).

Data Availability

Data from NHANES and MIDUS are publicly available from the Inter-university Consortium for Political and Social Research (<https://www.icpsr.umich.edu/web/pages/>). Data from the clinical trials described above are available upon reasonable request to the last author.

Conflicts of Interest

None declared.

Author Contributions

KB: Data curation, project administration, writing - original draft, writing - review and editing

GMS: Writing - original draft, writing - review and editing

IB: Writing - original draft, writing - review and editing

ALF: Data curation, formal analysis, writing - original draft, writing - review and editing

RS: Writing - original draft, writing - review and editing

LK: Writing - review and editing

MLB: Conceptualization, investigation, supervision, writing - review and editing

JAM: Conceptualization, data curation, methodology, formal analysis, writing - original draft, writing - review and editing

DA: Conceptualization, funding acquisition, methodology, supervision, writing - original draft, writing - review and editing

Abbreviations

BMI: Body mass index

BDI: Beck Depression Inventory

CES-D: Center for Epidemiologic Studies Depression Scale

CVD: Cardiovascular disease

INCOM: Iowa–Netherlands Comparison Orientation Measure

MIDUS: Midlife in the United States Study

MIMS: Monitor-Independent Measurement Summary units

MLM: Multilevel model

MVPA: moderate-to-vigorous physical activity

NHANES: National Health and Nutrition Examination Survey

PA: physical activity

PHQ-9: Patient Health Questionnaire

TSRQ: Treatment Self-Regulation Questionnaire

WALI: Weight and Lifestyle Inventory

References

1. Brim OG, Ryff CD, Kessler RC. *How Healthy Are We?: A National Study of Well-Being at Midlife*. University of Chicago Press; 2019.
2. Rodgers JL, Jones J, Bolleddu SI, et al. Cardiovascular Risks Associated with Gender and Aging. *J Cardiovasc Dev Dis*. 2019;6(2):19. doi:10.3390/jcdd6020019 PMID: 31035613
3. Matthews KA, Crawford SL, Chae CU, et al. Are Changes in Cardiovascular Disease Risk Factors in Midlife Women Due to Chronological Aging or to the Menopausal Transition? *J Am Coll Cardiol*. 2009;54(25):2366-2373. doi:10.1016/j.jacc.2009.10.009 PMID: 20082925
4. Appelman Y, van Rijn BB, ten Haaf ME, Boersma E, Peters SAE. Sex differences in cardiovascular risk factors and disease prevention. *Atherosclerosis*. 2015;241(1):211-218. doi:10.1016/j.atherosclerosis.2015.01.027 PMID: 25670232
5. Karvinen S, Jergenson MJ, Hyvärinen M, et al. Menopausal Status and Physical Activity Are Independently Associated With Cardiovascular Risk Factors of Healthy Middle-Aged Women: Cross-Sectional and Longitudinal Evidence. *Front Endocrinol*. 2019;10. doi:10.3389/fendo.2019.00589 PMID: 31543865
6. Kapoor E, Collazo-Clavell ML, Faubion SS. Weight Gain in Women at Midlife: A Concise Review of the Pathophysiology and Strategies for Management. *Mayo Clin Proc*. 2017;92(10):1552-1558. doi:10.1016/j.mayocp.2017.08.004 PMID: 28982486
7. Piercy KL, Troiano RP. Physical Activity Guidelines for Americans From the US Department of Health and Human Services: Cardiovascular Benefits and Recommendations. *Circ Cardiovasc Qual Outcomes*. 2018;11(11):e005263. doi:10.1161/CIRCOUTCOMES.118.005263 PMID: 30571339
8. Shiroma EJ, Lee IM. Physical Activity and Cardiovascular Health: Lessons Learned From Epidemiological Studies Across Age, Gender, and Race/Ethnicity. *Circulation*. 2010;122(7):743-752. doi:10.1161/CIRCULATIONAHA.109.914721 PMID: 20713909
9. Centers for Medicare & Medicaid Services. National Health Expenditure Data. 2019. Accessed January 8, 2023. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Age-and-Gender>
10. Waters LA, Galichet B, Owen N, Eakin E. Who Participates in Physical Activity Intervention Trials? *J Phys Act Health*. 2011;8(1):85-103. doi:10.1123/jpah.8.1.85 PMID: 21297189
11. Cooke R, Jones A. Recruiting adult participants to physical activity intervention studies using sport: a systematic review. *BMJ Open Sport Exerc Med*. 2017;3(1):e000231. doi:10.1136/bmjsem-2017-000231 PMID: 28761714
12. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: A systematic review and meta-analysis. *Soc Sci Med*. 2017;192:125-133.

doi:10.1016/j.socscimed.2017.09.021 PMID: 28965003

13. Hooker SP, Harmon B, Burroughs EL, Rheaume CE, Wilcox S. Exploring the feasibility of a physical activity intervention for midlife African American men. *Health Educ Res.* 2011;26(4):732-738. doi:10.1093/her/cyr034 PMID: 21597100
14. Mellor D, Connaughton C, McCabe MP, Tatangelo G. Better with age: A health promotion program for men at midlife. *Psychol Men Masculinities.* 2017;18(1):40-49. doi:10.1037/men0000037
15. Ballon-Landa E, Parsons JK. Nutrition, physical activity, and lifestyle factors in prostate cancer prevention. *Curr Opin Urol.* 2018;28(1):55-61. doi:10.1097/MOU.0000000000000460 PMID: 29049045
16. Griffith DM, Gunter K, Allen JO. Male Gender Role Strain as a Barrier to African American Men's Physical Activity. *Health Educ Behav.* 2011;38(5):482-491. doi:10.1177/1090198110383660 PMID: 21632436
17. Arigo D, Romano KA, Pasko K, et al. A scoping review of behavior change techniques used to promote physical activity among women in midlife. *Front Psychol.* 2022;13. doi:10.3389/fpsyg.2022.855749 PMID: 36211932
18. Grossman JA, Arigo D, Bachman JL. Meaningful weight loss in obese postmenopausal women: a pilot study of high-intensity interval training and wearable technology. *Menopause.* 2018;25(4):465-470. doi:10.1097/GME.0000000000001013 PMID: 29088015
19. Katzmarzyk PT, Lee IM, Martin CK, Blair SN. Epidemiology of Physical Activity and Exercise Training in the United States. *Prog Cardiovasc Dis.* 2017;60(1):3-10. doi:10.1016/j.pcad.2017.01.004 PMID: 28089610
20. Nomura T, Kawae T, Kataoka H, Ikeda Y. Aging, physical activity, and diabetic complications related to loss of muscle strength in patients with type 2 diabetes. *Phys Ther Res.* 2018;21(2):33-38. doi:10.1298/ptr.R0002 PMID: 30697507
21. Parvathaneni K, Surapaneni A, Ballew SH, et al. Association Between Midlife Physical Activity and Incident Kidney Disease: The Atherosclerosis Risk in Communities (ARIC) Study. *Am J Kidney Dis.* 2021;77(1):74-81. doi:10.1053/j.ajkd.2020.07.020 PMID: 32971191
22. Steene-Johannessen J, Anderssen SA, Van Der Ploeg HP, et al. Are Self-report Measures Able to Define Individuals as Physically Active or Inactive? *Med Sci Sports Exerc.* 2016;48(2):235-244. doi:10.1249/MSS.0000000000000760 PMID: 26322556
23. Jaeschke L, Steinbrecher A, Jeran S, Konigorski S, Pischon T. Variability and reliability study of overall physical activity and activity intensity levels using 24 h-accelerometry-assessed data. *BMC Public Health.* 2018;18(1):530. doi:10.1186/s12889-018-5415-8 PMID: 29678152
24. Craig CL, Tudor-Locke C, Cragg S, Cameron C. Process and Treatment of Pedometer Data Collection for Youth: The Canadian Physical Activity Levels among Youth Study. *Med Sci Sports Exerc.* 2010;42(3):430-435. doi:10.1249/MSS.0b013e3181b67544 PMID: 19952820

25. Baumann S, Groß S, Voigt L, et al. Pitfalls in accelerometer-based measurement of physical activity: The presence of reactivity in an adult population. *Scand J Med Sci Sports*. 2018;28(3):1056-1063. doi:10.1111/sms.12977 PMID: 28921747
26. Scott JJ, Morgan PJ, Plotnikoff RC, Trost SG, Lubans DR. Adolescent pedometer protocols: examining reactivity, tampering and participants' perceptions. *J Sports Sci*. 2014;32(2):183-190. doi:10.1080/02640414.2013.815361 PMID: 24015965
27. French DP, Sutton S. Reactivity of measurement in health psychology: How much of a problem is it? What can be done about it? *Br J Health Psychol*. 2010;15(3):453-468. doi:10.1348/135910710X492341 PMID: 20205982
28. König LM, Allmeta A, Christlein N, Van Emmenis M, Sutton S. A systematic review and meta-analysis of studies of reactivity to digital in-the-moment measurement of health behaviour. *Health Psychol Rev*. 2022;16(4):551-575. doi:10.1080/17437199.2022.2047096 PMID: 35264084
29. Zhu X, Haegele JA. Reactivity to Accelerometer Measurement of Children With Visual Impairments and Their Family Members. *Adapt Phys Act Q*. 2019;36(4):492-500. doi:10.1123/apaq.2019-0040 PMID: 31509801
30. Motl RW, McAuley E, Dlugonski D. Reactivity in baseline accelerometer data from a physical activity behavioral intervention. *Health Psychol*. 2012;31(2):172-175. doi:10.1037/a0025965 PMID: 22023436
31. Clemes SA, Deans NK. Presence and Duration of Reactivity to Pedometers in Adults. *Med Sci Sports Exerc*. 2012;44(6):1097. doi:10.1249/MSS.0b013e318242a377 PMID: 22595985
32. Conner T, Lehman B. Getting Started: Launching a Study in Daily Life. In: *Handbook of Research Methods For Studying Daily Life*. ; 2012:89-107.
33. Davis RE, Loprinzi PD. Examination of Accelerometer Reactivity Among a Population Sample of Children, Adolescents, and Adults. *J Phys Act Health*. 2016;13(12):1325-1332. doi:10.1123/jpah.2015-0703
34. Niemelä M, Kangas M, Farrahi V, et al. Intensity and temporal patterns of physical activity and cardiovascular disease risk in midlife. *Prev Med*. 2019;124:33-41. doi:10.1016/j.ypmed.2019.04.023
35. Vaccarezza M, Papa V, Milani D, et al. Sex/Gender-Specific Imbalance in CVD: Could Physical Activity Help to Improve Clinical Outcome Targeting CVD Molecular Mechanisms in Women? *Int J Mol Sci*. 2020;21(4):1477. doi:10.3390/ijms21041477
36. Klenk J, Peter RS, Rapp K, et al. Lazy Sundays: role of day of the week and reactivity on objectively measured physical activity in older people. *Eur Rev Aging Phys Act*. 2019;16(1):18. doi:10.1186/s11556-019-0226-1
37. Hilgenkamp T, Van Wijck R, Evenhuis H. Measuring Physical Activity with Pedometers in Older Adults with Intellectual Disability: Reactivity and Number of Days. *Intellect Dev Disabil*. 2012;50(4):343-351.

38. French DP, Miles LM, Elbourne D, et al. Reducing bias in trials due to reactions to measurement: experts produced recommendations informed by evidence. *J Clin Epidemiol*. 2021;139:130-139. doi:10.1016/j.jclinepi.2021.06.028
39. König LM, Pasko K, Baga K, Harsora R, Arigo D. Isolating the role of researcher observation on reactivity to the measurement of physical activity. *Appl Psychol Health Well-Being*. 2024 (in press).
40. Simon GE, Ludman EJ, Linde JA, et al. Association between obesity and depression in middle-aged women. *Gen Hosp Psychiatry*. 2008;30(1):32-39. doi:10.1016/j.genhosppsych.2007.09.001
41. Hansen BH, Holme I, Anderssen SA, Kolle E. Patterns of Objectively Measured Physical Activity in Normal Weight, Overweight, and Obese Individuals (20–85 Years): A Cross-Sectional Study. *PLOS ONE*. 2013;8(1):e53044. doi:10.1371/journal.pone.0053044
42. Lachman ME, Lipsitz L, Lubben J, Castaneda-Sceppa C, Jette AM. When Adults Don't Exercise: Behavioral Strategies to Increase Physical Activity in Sedentary Middle-Aged and Older Adults. *Innov Aging*. 2018;2(1):igy007. doi:10.1093/geroni/igy007
43. Gibbons FX, Buunk BP. Individual differences in social comparison: Development of a scale of social comparison orientation. *J Pers Soc Psychol*. 1999;76(1):129-142. doi:10.1037/0022-3514.76.1.129
44. Sreedhara M, Silfee VJ, Rosal MC, Waring ME, Lemon SC. Does provider advice to increase physical activity differ by activity level among US adults with cardiovascular disease risk factors? *Fam Pract*. 2018;35(4):420-425. doi:10.1093/fampra/cmz140
45. Hebert JR, Ma Y, Clemow L, et al. Gender Differences in Social Desirability and Social Approval Bias in Dietary Self-report. *Am J Epidemiol*. 1997;146(12):1046-1055. doi:10.1093/oxfordjournals.aje.a009233
46. Steinhardt MA, Dishman RK. Reliability and Validity of Expected Outcomes and Barriers for Habitual Physical Activity: *J Occup Environ Med*. 1989;31(6):536-546. doi:10.1097/00043764-198906000-00011
47. Barta WD, Tennen H, Litt MD. Measurement reactivity in diary research. In: *Handbook of Research Methods for Studying Daily Life*. The Guilford Press; 2012:108-123.
48. Fuller D, Colwell E, Low J, et al. Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. *JMIR MHealth UHealth*. 2020;8(9):e18694. doi:10.2196/18694
49. Kanejima Y, Kitamura M, Izawa KP. Self-monitoring to increase physical activity in patients with cardiovascular disease: a systematic review and meta-analysis. *Aging Clin Exp Res*. 2019;31(2):163-173. doi:10.1007/s40520-018-0960-7
50. Hofer SM, Piccinin AM. Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychol Methods*. 2009;14(2):150-164. doi:10.1037/a0015566

51. Jones C. Archival data: Advantages and disadvantages for research in psychology. *Soc Personal Psychol Compass*. 2010;4(11):1008-1017. doi:10.1111/j.1751-9004.2010.00317.x
52. NHANES Questionnaires, Datasets, and Related Documentation. Accessed October 3, 2024. <https://www.cdc.gov/nchs/nhanes/>
53. Butryn ML, Forman EM, Lowe MR, Gorin AA, Zhang F, Schaumberg K. Efficacy of environmental and acceptance-based enhancements to behavioral weight loss treatment: The ENACT trial. *Obesity*. 2017;25(5):866-872. doi:10.1002/oby.21813
54. Butryn ML, Godfrey KM, Call CC, Forman EM, Zhang F, Volpe SL. Promotion of physical activity during weight loss maintenance: A randomized controlled trial. *Health Psychol*. 2021;40(3):178-187. doi:10.1037/hea0001043
55. Butryn ML, Martinelli MK, Crane NT, et al. Counselor Surveillance of Digital Self-Monitoring Data: A Pilot Randomized Controlled Trial. *Obesity*. 2020;28(12):2339-2346. doi:10.1002/oby.23015
56. Miller NA, Ehmann MM, Hagerman CJ, et al. Sharing digital self-monitoring data with others to enhance long-term weight loss: A randomized controlled trial. *Contemp Clin Trials*. 2023;129:107201. doi:10.1016/j.cct.2023.107201
57. John D, Tang Q, Albinali F, Intille S. An Open-Source Monitor-Independent Movement Summary for Accelerometer Data Processing. *J Meas Phys Behav*. 2019;2(4):268-281. doi:10.1123/jmpb.2018-0068
58. Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. *Psychiatr Ann*. 2002;32(9):509-515.
59. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Appl Psychol Meas*. 1977;1(3):385-401. doi:10.1177/014662167700100306
60. Beck AT, Steer RA, Brown G. Beck depression inventory-II (BDI-II)[Database record]. *APA PsycTests*. 1996;10.
61. Wadden TA, Foster GD. Weight and Lifestyle Inventory (WALI). *Obesity*. 2006;14(S3):99S-118S. doi:10.1038/oby.2006.289
62. Levesque CS, Williams GC, Elliot D, Pickering MA, Bodenhamer B, Finley PJ. Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. *Health Educ Res*. 2006;22(5):691-702. doi:10.1093/her/cyl148
63. Mogle J, Hill NL, Turner JR. Individual differences and features of self-reported memory lapses as risk factors for Alzheimer disease among adults aged 50 years and older: Protocol for a coordinated analysis across two longitudinal data sets. *JMIR Res Protoc*. 2021;10(5):e25233. doi:10.2196/25233
64. Hill NL, Bhargava S, Bratlee-Whitaker E, Turner JR, Brown MJ, Mogle J. Longitudinal relationships between subjective cognitive decline and objective memory: Depressive symptoms

- mediate between-person associations. *J Alzheimers Dis.* 2021;83(4):1623-1636. doi:10.3233/JAD-210230
65. Arigo D, König LM. Examining reactivity to the measurement of physical activity and sedentary behavior among women in midlife with elevated risk for cardiovascular disease. *Psychol Health*. Published online April 12, 2022:1-17. doi:10.1080/08870446.2022.2055024
66. Arigo D, Pasko K, Mogle JA. Daily relations between social perceptions and physical activity among college women. *Psychol Sport Exerc.* 2020;47:101528. doi:10.1016/j.psychsport.2019.04.018
67. Arigo D, Mogle JA, Smyth JM. Relations between social comparisons and physical activity among women in midlife with elevated risk for cardiovascular disease: an ecological momentary assessment study. *J Behav Med.* 2021;44(5):579-590. doi:10.1007/s10865-021-00229-7
68. Arigo D, Roberts SR, Butryn ML. Social comparisons between group members during behavioural weight loss treatment: comparison direction, scale, and associations with weight loss maintenance. *Psychol Health.* 2023;38(4):429-444. doi:10.1080/08870446.2021.1967953
69. Maher JP, Arigo D, Baga K, et al. Measurement Reactivity in Ecological Momentary Assessment Studies of Movement-Related Behaviors. *J Meas Phys Behav.* 2024;7(1):jmpb.2023-0035. doi:10.1123/jmpb.2023-0035
70. Lee IM, Moore CC, Evenson KR. Maximizing the Utility and Comparability of Accelerometer Data From Large-Scale Epidemiologic Studies. Published online January 11, 2023. doi:10.1123/jmpb.2022-0035
71. Hox J, Moerbeek M, Schoot R van de. *Multilevel Analysis: Techniques and Applications*. 3rd ed. Routledge; 2017. doi:10.4324/9781315650982
72. Hoffman L. *Longitudinal Analysis: Modeling Within-Person Fluctuation and Change*. 1st ed. Routledge; 2015. doi:10.4324/9781315744094
73. French DP, Miles LM, Elbourne D, et al. Reducing bias in trials from reactions to measurement: the MERIT study including developmental work and expert workshop. *Health Technol Assess.* 2021;25(55):1-72. doi:10.3310/hta25550
74. Kanejima Y, Kitamura M, Izawa KP. Self-monitoring to increase physical activity in patients with cardiovascular disease: a systematic review and meta-analysis. *Aging Clin Exp Res.* 2019;31(2):163-173. doi:10.1007/s40520-018-0960-7
75. Arigo D, Brown MM, Pasko K, Ainsworth MC, Travers L, Gupta A, Downs DS, Smyth JM. Rationale and design of the Women's Health And Daily Experiences Project: Protocol for an ecological momentary assessment study to identify real-time predictors of midlife women's physical activity. *JMIR Res Protoc.* 2020;9(10):e19044. doi:10.2196/19044
76. Infurna FJ, Gerstorf D, Lachman ME. Midlife in the 2020s: Opportunities and challenges. *Am Psychol.* 2020;75(4):470-485. doi:10.1037/amp0000591
77. Anderson A, Getz K, Connor S, Getz N. Trends in global public and patient perceptions of

clinical research. *Appl Clin Trials*. 2018;27(11):18-23.

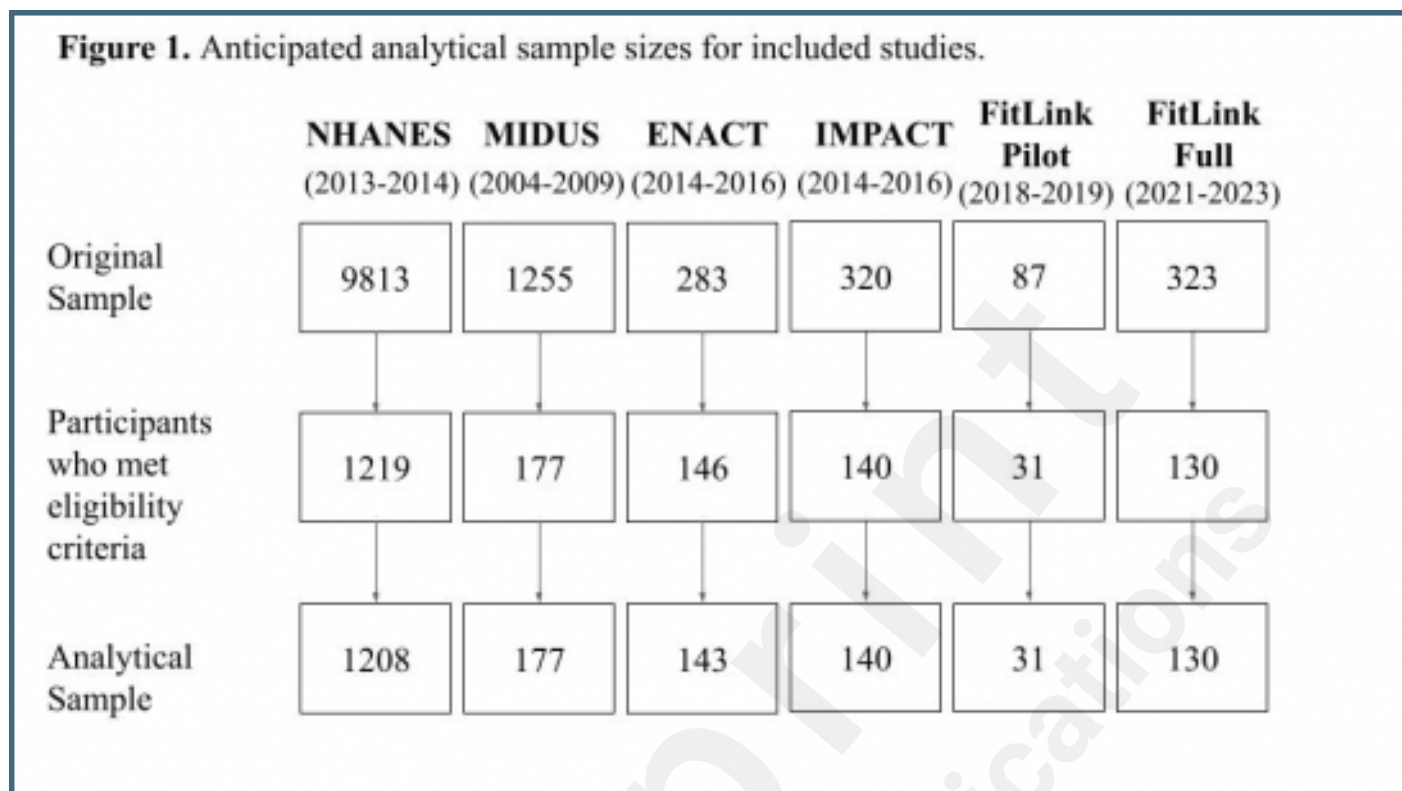
78. Gabel M, Bollinger RM, Knox M, Coble DW, Grill JD, Edwards DF, Stark SL, Lingler JH. Perceptions of research burden and retention among participants in ADRC cohorts. *Alzheimer Dis Assoc Disord*. 2022;36(4):281-287. doi:10.1097/WAD.0000000000000514



Supplementary Files

Figures

Anticipated analytical sample sizes for included studies.



Multimedia Appendixes

Peer reviewer comments from NIH grant review.

URL: <http://asset.jmir.pub/assets/87bd6dfef60d01675ac108309b7fcb7e.pdf>

